GenCore version 5.1.9

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 20:07:43 ; Search time 2712 Seconds
(without alignments)
9431.773 Million cell updates/sec

Title: SEQ1-47502C
Perfect score: 399.6
Sequence: 1 ccaggtactcagccatgtgc.....tattgcaggagaccaaaaag 400

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : GenEmbl:*
1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_sts:*
8: gb_sy:*
9: gb_un:*
10: gb_vi:*
11: gb_ov:*
12: gb_htg:*
13: gb_in:*
14: gb_om:*
15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	398	99.6	187064	5	AC011476	AC011476 Homo sapi
	2	231	57.8	204340	12	AC019238	AC019238 Homo sapi
c	3	161.2	40.3	184822	12	AC146499	AC146499 Aotus nan
	4	160.6	40.2	232406	12	AC146898	AC146898 Pongo pyg
	5	160	40.0	97676	5	AC005497	AC005497 Homo sapi
c	6	159.8	40:0	107868	5	AL445686	AL445686 Human DNA
	7	159.8	40.0	171480	5	AC016542	AC016542 Homo sapi
	8	159.4	39.9	97916	5	HS408N23	Z98048 Human DNA s
c	9	158.8	39.7	195616	12	AC090344	AC090344 Homo sapi
c	10	158.8	39.7	197156	12	AC090345	AC090345 Homo sapi
	11	158.8	39.7	197856	5	AP000777	AP000777 Homo sapi
c	12	158.6	39.7	68990	12	AC087662	AC087662 Homo sapi
	13	158.4	39.6	96593	2	DD164421	DD164421 NOVEL COM
	14	158.4	39.6	96593	2	AX695641	AX695641 Sequence
c	15	158.4	39.6	112659	5	AC010677	AC010677 Homo sapi
	16	158.2	39.6	155344	5	AC026407	AC026407 Homo sapi
	17	157.8	39.5	170154	5	AC040168	AC040168 Homo sapi
	18	157.8	39.5	186747	12	AC018605	AC018605 Homo sapi
	19	157.8	39.5	198575	5	AC007495	AC007495 Homo sapi
	20	157.8	39.5	224187	5	AL732374	AL732374 Human DNA
c	21	157.6	39.4	169908	12	AC150910	AC150910 Pan trogl
	22	157.4	39.4	174097	5	AC069513	AC069513 Homo sapi
	23	157.4	39.4	209844	5	AC011495	AC011495 Homo sapi
	24	157.2	39.3	173556	12	AC074388	AC074388 Homo sapi
c	25	157	39.3	84001	5	AL162739	AL162739 Human DNA
c	26	156.8	39.2	158330	5	AC025599	AC025599 Homo sapi
c	27	156.8	39.2	195070	12	AC068995	AC068995 Homo sapi
c	28	156.6	39.2	111372	12	AL353145	AL353145 Homo sapi
	29	156.6	39.2	150934	12	AC022252	AC022252 Homo sapi
c	30	156.6	39.2	163521	12	AC108040	AC108040 Homo sapi
	31	156.6	39.2	196773	5	AC114480	AC114480 Homo sapi
c	32	156.4	39.1	102008	12	AC016552	AC016552 Homo sapi
c	33	156.4	39.1	244525	5	AC008499	AC008499 Homo sapi
	34	155.8	39.0	88848	5	AC107377	AC107377 Homo sapi
c	35	155.8	39.0	182725	12	AC172796	AC172796 Callicebu
	36	155.8	39.0	216387	5	DJ534K4	AF109907 Homo sapi
c	37	155.6	38.9	224187	5	AL732374	AL732374 Human DNA
	38	155.4	38.9	122961	5	HSJ193N13	AL078600 Human DNA
	39	155.4	38.9	126462	5	AC004876	AC004876 Homo sapi
c	40	155.4	38.9	132492	5	AC007616	AC007616 Homo sapi
c	41	155.4	38.9	145679	12	AC027249	AC027249 Homo sapi
	42	155.4	38.9	183444	5	AP001024	AP001024 Homo sapi
c	43	155.4	38.9	202138	12	AC018423	AC018423 Homo sapi
	44	155.4	38.9	204493	5	AC099489	AC099489 Homo sapi
	45	155.4	38.9	205463	5	AL355388	AL355388 Human DNA
	46	155.2	38.8	116130	5	AC026423	AC026423 Homo sapi
c	47	155.2	38.8	177640	5	AC079456	AC079456 Homo sapi
c	48	155.2	38.8	178079	12	AC068929	AC068929 Homo sapi
	49	155	38.8	68465	5	AL162384	AL162384 Human DNA
	50	155	38.8	73433	5	AC068889	AC068889 Homo sapi
	51	155	38.8	168684	12	AC023985	AC023985 Homo sapi
	52	155	38.8	171811	5	AC011739	AC011739 Homo sapi
c	53	155	38.8	196216	5	AC099343	AC099343 Homo sapi
	54	154.8	38.7	98832	5	AL137127	AL137127 Human DNA
c	55	154.8	38.7	152757	5	AC018475	AC018475 Homo sapi
	56	154.6	38.7	174361	5	AL158044	AL158044 Human DNA
	57	154.4	38.6	126525	5	HSJ1077B9	AL096840 Human DNA
c	58	154.4	38.6	159273	12	AC025001	AC025001 Homo sapi
	59	154.4	38.6	169714	12	AC138876	AC138876 Homo sapi
	60	154.4	38.6	182892	5	AC034244	AC034244 Homo sapi
c	61	154.4	38.6	200518	12	AC151041	AC151041 Callithri
	62	154.2	38.6	638	7	BV472608	BV472608 G591P6278
	63	154.2	38.6	74996	5	AP001986	AP001986 Homo sapi
c	64	154.2	38.6	122748	5	HS742C19	AL031846 Human DNA
	65	154.2	38.6	161166	12	AC087673	AC087673 Homo sapi
c	66	154	38.5	20425	12	BX284653_5	Continuation (6 of
c	67	154	38.5	110000	12	BX284653_4	Continuation (5 of
	68	154	38.5	119389	5	AL450328	AL450328 Human DNA
c	69	154	38.5	154612	5	AL365504	AL365504 Human DNA
c	70	154	38.5	172095	12	AC026834	AC026834 Homo sapi
c	71	154	38.5	176813	12	AC107895	AC107895 Homo sapi
	72	154	38.5	198821	5	AC120057	AC120057 Homo sapi
c	73	154	38.5	206773	12	AC010932	AC010932 Homo sapi
	74	153.8	38.5	115246	12	AC161015	AC161015 Pan trogl
c	75	153.8	38.5	121295	5	AL442071	AL442071 Human DNA
c	76	153.8	38.5	154616	5	AC004846	AC004846 Homo sapi
	77	153.8	38.5	173333	12	AC025282	AC025282 Homo sapi
c	78	153.8	38.5	175347	12	AC009099	AC009099 Homo sapi
	79	153.8	38.5	203046	12	AC006342	AC006342 Homo sapi
	80	153.8	38.5	206258	12	AC112775	AC112775 Homo sapi
c	81	153.8	38.5	233418	12	AC109128	AC109128 Homo sapi
	82	153.6	38.4	139063	12	AC133563	AC133563 Homo sapi
c	83	153.6	38.4	146208	5	AL354776	AL354776 Human DNA
	84	153.6	38.4	167101	5	AC002094	AC002094 Genomic s
c	85	153.6	38.4	174724	12	AC140096	AC140096 Pan trogl
	86	153.6	38.4	176181	5	AC008155	AC008155 Homo sapi
c	87	153.6	38.4	184684	12	AC148258	AC148258 Callithri
c	88	153.4	38.4	67396	5	AC008848	AC008848 Homo sapi
c	89	153.4	38.4	141605	5	AL355353	AL355353 Human DNA
	90	153.4	38.4	161251	12	AC147284	AC147284 Pan trogl
	91	153.4	38.4	166697	12	AC021103	AC021103 Homo sapi

c	92	153.4	38.4	203480	12	AC146136 AC146136 Pan trogl
c	93	153.2	38.3	80436	5	AC114759 AC114759 Homo sapi
	94	153.2	38.3	114420	5	AL158076 AL158076 Human DNA
	95	153.2	38.3	114784	5	HS1189B24 AL030996 Human DNA
	96	153.2	38.3	128729	5	AC137590 AC137590 Homo sapi
c	97	153.2	38.3	183411	12	AC137495 AC137495 Homo sapi
	98	153.2	38.3	184453	5	AC025593 AC025593 Homo sapi
	99	153.2	38.3	270567	12	AC138086 AC138086 Homo sapi
c	100	153	38.3	34335	5	AC112711 AC112711 Homo sapi
	101	153	38.3	170758	5	AC004965 AC004965 Homo sapi
	102	152.8	38.2	102562	12	AC109516 AC109516 Homo sapi
	103	152.8	38.2	166942	2	CS247571 CS247571 Sequence
	104	152.8	38.2	166942	5	HS1054A22 AL031651 Human DNA
c	105	152.8	38.2	167509	12	AC012107 AC012107 Homo sapi
c	106	152.6	38.2	67165	5	AL139089 AL139089 Human DNA
	107	152.6	38.2	128479	5	AL596266 AL596266 Human DNA
c	108	152.6	38.2	138350	12	AC010801 AC010801 Homo sapi
	109	152.6	38.2	144588	12	AC012239 AC012239 Homo sapi
c	110	152.6	38.2	146741	5	AC011452 AC011452 Homo sapi
c	111	152.6	38.2	168931	12	AC144998 AC144998 Pan trogl
	112	152.6	38.2	169074	12	AC148830 AC148830 Pan trogl
	113	152.6	38.2	175999	12	AC021154 AC021154 Homo sapi
c	114	152.6	38.2	190430	12	AC118987 AC118987 Pan trogl
	115	152.4	38.1	576	7	BV198237 BV198237 sqnm19689
c	116	152.4	38.1	63441	5	AL356502 AL356502 Human DNA
c	117	152.4	38.1	126382	12	AC133392 AC133392 Pan trogl
c	118	152.4	38.1	142170	5	AC087362 AC087362 Homo sapi
	119	152.4	38.1	143847	5	AF205589 AF205589 Homo sapi
	120	152.4	38.1	160864	12	AC021993 AC021993 Homo sapi
	121	152.4	38.1	162835	5	AC087623 AC087623 Homo sapi
	122	152.4	38.1	163494	5	AC079865 AC079865 Homo sapi
	123	152.4	38.1	166237	5	AL139379 AL139379 Human DNA
c	124	152.4	38.1	170338	12	AF214633 AF214633 Homo sapi
	125	152.4	38.1	174562	12	AC005848 AC005848 Homo sapi
	126	152.4	38.1	194646	12	AC090129 AC090129 Homo sapi
c	127	152.4	38.1	205577	12	AC011237 AC011237 Homo sapi
	128	152.2	38.1	615	7	BV504898 BV504898 qpq33e03.
c	129	152.2	38.1	51242	5	BX120002 BX120002 Human DNA
c	130	152.2	38.1	149481	5	AC013472 AC013472 Homo sapi
c	131	152.2	38.1	155735	5	AC007683 AC007683 Homo sapi
	132	152.2	38.1	159726	5	AC125257 AC125257 Homo sapi
	133	152.2	38.1	168017	5	BS000238 BS000238 Pan trogl
c	134	152.2	38.1	169604	5	AC083884 AC083884 Homo sapi
c	135	152.2	38.1	175303	12	AC024721 AC024721 Homo sapi
	136	152.2	38.1	178345	5	AC091172 AC091172 Homo sapi
	137	152.2	38.1	190477	5	PTB109N16 AL954242 Pan trogl
c	138	152.2	38.1	196735	12	AC012192 AC012192 Homo sapi
c	139	152.2	38.1	197503	12	AC165190 AC165190 Colobus g
	140	152	38.0	147118	5	AC007513 AC007513 Homo sapi
c	141	152	38.0	174424	2	AX335950 AX335950 Sequence
c	142	152	38.0	179676	12	AC129889 AC129889 Papio anu
c	143	152	38.0	181343	5	U52112 U52112 Homo sapien
	144	152	38.0	184989	12	AC113392 AC113392 Homo sapi
c	145	152	38.0	223606	12	AC013479 AC013479 Homo sapi
c	146	152	38.0	225370	5	AL445222 AL445222 Human DNA
c	147	151.8	38.0	169323	12	AL161906 AL161906 Homo sapi
c	148	151.8	38.0	169515	5	AP002777 AP002777 Homo sapi
c	149	151.6	37.9	743	7	BV642952 BV642952 S216P6376
c	150	151.6	37.9	40536	12	AC110593 AC110593 Homo sapi

ALIGNMENTS

```
RESULT 1
AC011476
LOCUS       AC011476               187064 bp    DNA     linear   PRI 16-AUG-2002
DEFINITION  Homo sapiens chromosome 19 clone CTC-550B14, complete sequence.
ACCESSION   AC011476
VERSION     AC011476.8  GI:22267569
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 187064)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 187064)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 187064)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-AUG-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Aug 16, 2002 this sequence version replaced gi:15022008.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.3.
            NOTE: Transposon sequencing failed to verify the number of repeat
            copies 67315-69359. Unsure number of repeat copies 67315-69359.
            NOTE:  Shatter libraries failed to resolve dinucleotide repeat
            region 171590-171722.  Unsure number of repeat copies
            171590-171722.  Forced join 171695.
FEATURES             Location/Qualifiers
     source          1. .187064
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTC-550B14"
     misc_feature    67315. .69359
                     /note="NOTE: Transposon sequencing failed to verify the
                     number of repeat copies 67315-69359. Unsure number of
                     repeat copies 67315-69359."
     misc_feature    171590. .171722
                     /note="NOTE:  Shatter libraries failed to resolve
                     dinucleotide repeat region 171590-171722.  Unsure number
                     of repeat copies 171590-171722.  Forced join 171695."
ORIGIN

  Query Match             99.6%;  Score 398;  DB 5;  Length 187064;
  Best Local Similarity   99.5%;  Pred. No. 5.2e-128;
  Matches  398;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCAGGTACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119458 CCAGGTACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGG 119517

Qy         61 CCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCG 120
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     119518 CCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCG 119577

Qy        121 GATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119578 GATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTA 119637

Qy        181 CTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119638 CTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGG 119697

Qy        241 AGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119698 AGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTG 119757

Qy        301 AGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 360
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db     119758 AGCAGAGTGAGACTCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 119817
```

```
Qy        361 TCCATCTGCGGTTCCCAGACTATTGCAGGAGACCAAAAAG 400
              |||||||||||||||||||||||||||||||||||||||
Db     119818 TCCATCTGCGGTTCCCAGACTATTGCAGGAGACCAAAAAG 119857


RESULT 2
AC019238
LOCUS       AC019238              204340 bp    DNA     linear   HTG 17-AUG-2000
DEFINITION  Homo sapiens chromosome 19 clone RP11-700B5, WORKING DRAFT
            SEQUENCE, 12 unordered pieces.
ACCESSION   AC019238
VERSION     AC019238.5  GI:9838316
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 204340)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 204340)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Aug 17, 2000 this sequence version replaced gi:9280808.

            ---------- Genome Center ----------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ---------- Project Information ----------
            Center project name: H_NH0700B05
            ---------- Summary Statistics  ----------
            Sequencing vector: M13; 84%
            Sequencing vector: plasmid; 16%
            Chemistry: Dye-primer ET; 84% of reads
            Chemistry: Dye-terminator Big Dye; 16% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 199676 bases at least Q40
            Consensus quality: 200928 bases at least Q30
            Consensus quality: 201749 bases at least Q20
            Insert size: 236000; agarose-fp
            Insert size: 204523; sum-of-contigs
            Quality coverage: 6.54 in Q20 bases; agarose-fp
            Quality coverage: 7.57 in Q20 bases; sum-of-contigs
            ----------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 12 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2752: contig of 2752 bp in length
            *     2753     2852: gap of unknown length
            *     2853     5901: contig of 3049 bp in length
            *     5902     6001: gap of unknown length
            *     6002    12965: contig of 6964 bp in length
            *    12966    13065: gap of unknown length
            *    13066    21018: contig of 7953 bp in length
            *    21019    21118: gap of unknown length
            *    21119    30399: contig of 9281 bp in length
            *    30400    30499: gap of unknown length
            *    30500    40958: contig of 10459 bp in length
            *    40959    41058: gap of unknown length
            *    41059    52180: contig of 11122 bp in length
            *    52181    52280: gap of unknown length
            *    52281    65275: contig of 12995 bp in length
            *    65276    65375: gap of unknown length
            *    65376    95161: contig of 29786 bp in length
            *    95162    95261: gap of unknown length
            *    95262   122538: contig of 27277 bp in length
            *   122539   122638: gap of unknown length
            *   122639   155778: contig of 33140 bp in length
            *   155779   155878: gap of unknown length
            *   155879   204340: contig of 48462 bp in length.
FEATURES             Location/Qualifiers
     source          1. .204340
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="RP11-700B5"
     misc_feature    1. .2752
                     /note="assembly_name:Contig16"
     gap             2753. .2852
                     /estimated_length=unknown
     misc_feature    2853. .5901
                     /note="assembly_name:Contig17"
     gap             5902. .6001
                     /estimated_length=unknown
     misc_feature    6002. .12965
                     /note="assembly_name:Contig18"
     gap             12966. .13065
                     /estimated_length=unknown
     misc_feature    13066. .21018
                     /note="assembly_name:Contig19"
     gap             21019. .21118
                     /estimated_length=unknown
     misc_feature    21119. .30399
                     /note="assembly_name:Contig20"
     gap             30400. .30499
                     /estimated_length=unknown
     misc_feature    30500. .40958
                     /note="assembly_name:Contig21"
     gap             40959. .41058
                     /estimated_length=unknown
     misc_feature    41059. .52180
                     /note="assembly_name:Contig22"
     gap             52181. .52280
                     /estimated_length=unknown
     misc_feature    52281. .65275
                     /note="assembly_name:Contig23"
     gap             65276. .65375
                     /estimated_length=unknown
     misc_feature    65376. .95161
                     /note="assembly_name:Contig24"
     gap             95162. .95261
                     /estimated_length=unknown
     misc_feature    95262. .122538
                     /note="assembly_name:Contig25"
     gap             122539. .122638
                     /estimated_length=unknown
     misc_feature    122639. .155778
                     /note="assembly_name:Contig26
                     clone_end:T7
                     vector_side:right"
     gap             155779. .155878
                     /estimated_length=unknown
     misc_feature    155879. .204340
                     /note="assembly_name:Contig27
                     clone_end:SP6
                     vector_side:left"
ORIGIN

  Query Match             57.8%;  Score 231;  DB 12;  Length 204340;
  Best Local Similarity   99.1%;  Pred. No. 8.9e-69;
  Matches  231;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        168 AACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTC 227
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65376 AACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTC 65435

Qy        228 CCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCAC 287
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65436 CCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCAC 65495

Qy        288 AGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
            |||||||||||||||| ||||||||||:||||||||||||||||||||||||||||||||||
Db      65496 AGATCACAGAGTGAGGAGAGTGAGACGCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 65555

Qy        348 CCATAAGACATTGTCCATCTGCGGTTCCCAGACTATTGCAGGAGACCAAAAAG 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65556 CCATAAGACATTGTCCATCTGCGGTTCCCAGACTATTGCAGGAGACCAAAAAG 65608


RESULT 3
AC146499/c
LOCUS       AC146499               184822 bp    DNA     linear   HTG 26-NOV-2003
DEFINITION  Aotus nancymaae clone CH258-450E24, WORKING DRAFT SEQUENCE, 2
            ordered pieces.
ACCESSION   AC146499
VERSION     AC146499.1  GI:34013545
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Aotus nancymaae (Ma's night monkey)
  ORGANISM  Aotus nancymaae
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Platyrrhini;
            Cebidae; Aotinae; Aotus.
REFERENCE   1  (bases 1 to 184822)
  AUTHORS   Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
            Peng,Z., Malinov,I. and Rubin,E.M.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 184822)
  AUTHORS   Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
            Peng,Z., Malinov,I. and Rubin,E.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-AUG-2003) Genome Sciences, Lawrence Berkeley National
            Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
REFERENCE   3  (bases 1 to 184822)
  AUTHORS   Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
            Peng,Z., Malinov,I. and Rubin,E.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-NOV-2003) Genome Sciences, Lawrence Berkeley National
            Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
COMMENT
            Sequence Produced by Berkeley PGA
            Web site: http://pga.lbl.gov
            Center Code: PGABERK
            Center Project Name: W033
            Bac Clone Name: CH258-450E24

            This sequence has been compared to sequences of other species
            using Vista (http://www-gsd.lbl.gov/VISTA). The results can be
            viewed at:
            http://pga.lbl.gov/cgi-bin/search_cvcgd?type=n&value=APOA1

            The order-orientation of the draft sequence was accomplished by
            using:
            Avid (http://baboon.math.berkeley.edu/mavid),
            Lagan (http://lagan.stanford.edu/) and paired end information.

            Funding agent: Programs for Genomic Applications (NHLBI)

            Summary Statistics:
            Sequencing vector: Plasmid; pUC18
            Chemistry: Dye-terminator Big Dye
            Assembly program: Phrap version 0.990329.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be provided.
            *       1    32135: contig of 32135 bp in length
            *   32136    32235: gap of unknown length
            *   32236   184822: contig of 152587 bp in length.
FEATURES            Location/Qualifiers
     source         1. .184822
                    /organism="Aotus nancymaae"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:37293"
                    /clone="CH258-450E24"
     gap            32136. .32235
                    /estimated_length=unknown
ORIGIN

  Query Match           40.3%;  Score 161.2;  DB 12;  Length 184822;
  Best Local Similarity  75.9%;  Pred. No. 4.7e-44;
  Matches  211;  Conservative    1;  Mismatches   64;  Indels    2;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| |||||||||||| ||| ||||||| ||||||||||||
Db     133786 ACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGTGGATCAC--GAGGTCAAGAGA 133729

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              |||||||||||||| |||||||||||| ||||||| ||||||||||||||| ||||
Db     133728 TCGAGACCATCCTGGTCAACATGGTGAAAACCCGTCTCTACTAAAAATACAAAAATTAGC 133669

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||||||||| |||||||| |||||||||| |||| | |||  | ||
Db     133668 TGGGCATGGTGGCATGCACCTGTAATCCCAGCTACTTGGGAGGCTGAGTCAGGAGAATTG 133609

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              || | |||   | |  |    |||      | | || |  |||||| ||:| ||||
Db     133608 CTTGAACCCAGGAGGCGGAGGTTGCACTCCAGCCTGGGTAACAAGAGTGAAACTCTGTCT 133549

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACAT 358
              |||||| || || |||||  ||| |||  |  | || |
Db     133548 CAAAAAAAAAAAAAAAAAGAAAAGAAAAGAATAAACTT 133511


RESULT 4
AC146898
LOCUS       AC146898               232406 bp    DNA     linear   HTG 03-DEC-2003
DEFINITION  Pongo pygmaeus clone CH253-50L14, WORKING DRAFT SEQUENCE, 4 ordered
            pieces.
ACCESSION   AC146898
VERSION     AC146898.3  GI:38638702
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 232406)
  AUTHORS   Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
            Coleman,H., Engle,J., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B., Idol,J.R.,
            Karlins,E., Kwong,P., Laric,P., Larson,S., Lee-Lin,S.-Q.,
            Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
            Maskeri,B., McDowell,J., Mullikin,J.C., Paguirigan,C., Pearson,R.,
            Portnoy,M.E., Prasad,A., Reddix-Dugue,N., Schandler,K.,
            Schueler,M.G., Shah,K., Sison,C., Stantripop,S., Thomas,J.W.,
            Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Young,A. and
            Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 232406)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
```

```
JOURNAL   Submitted (18-OCT-2003) NIH Intramural Sequencing Center, 8717
          Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE 3  (bases 1 to 232406)
  AUTHORS Green,E.D.
  TITLE   Direct Submission
  JOURNAL Submitted (03-DEC-2003) NIH Intramural Sequencing Center, 8717
          Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT   On Dec 3, 2003 this sequence version replaced gi:38016037.
          -------------- Genome Center
               Center: NIH Intramural Sequencing Center
               Center code: NISC
               Web site: http://www.nisc.nih.gov
               Contact: nisc_zoo@nhgri.nih.gov
          -------------- Project Information
               Center project name: ely
               Center clone name: 050L14

          The sequence data in this record represents an 'enhanced'
          version of a Phase 2 submission. Specifically, the indicated
          order and orientation of each sequence contig has been
          established using one or more of the following: read-pair
          data from individual subclones, overlaps with neighboring
          clones, alignment with available reference sequence (e.g.,
          human), and/or confirmation by PCR testing. In addition,
          the sequence assembly is based on at least 8X average
          coverage in Q20 bases and has been reviewed to rule out
          gross misassemblies, the low-quality ends of sequence
          contigs have been trimmed away, and each base is associated
          with a Phrap-derived quality score.
          -------------- Summary Statistics
               Sequencing vector: plasmid; n/a; 100% of reads
               Chemistry: Dye-terminator Big Dye; 100% of reads
               Assembly program: Phrap; version 0.990319
               Consensus quality: 231482 bases at least Q40
               Consensus quality: 231897 bases at least Q30
               Consensus quality: 232064 bases at least Q20
               Insert size: 245000; agarose-fp
               Insert size: 232106; sum-of-contigs
               Quality coverage: 21.15x in Q20 bases; agarose-fp
               Quality coverage: 22.32x in Q20 bases; sum-of-contigs
          --------------.
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 4 contigs. Gaps between the contigs
          * are represented as runs of N. The order of the pieces
          * is believed to be correct as given, however the sizes
          * of the gaps between them are based on estimates that have
          * provided by the submittor.
          * This sequence will be replaced
          * by the finished sequence as soon as it is available and
          * the accession number will be preserved.
          *      1    80408: contig of 80408 bp in length
          *  80409    80508: gap of unknown length
          *  80509   101957: contig of 21449 bp in length
          * 101958   102057: gap of unknown length
          * 102058   206149: contig of 104092 bp in length
          * 206150   206249: gap of unknown length
          * 206250   232406: contig of 26157 bp in length.
FEATURES             Location/Qualifiers
     source          1. .232406
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9600"
                     /clone="CH253-50L14"
                     /clone_lib="CH253"
     misc_feature    1. .80408
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:left"
     gap             80409. .80508
                     /estimated_length=unknown
     misc_feature    80509. .101957
                     /note="assembly_fragment"
     gap             101958. .102057
                     /estimated_length=unknown
     misc_feature    102058. .206149
                     /note="assembly_fragment"
     gap             206150. .206249
                     /estimated_length=unknown
     misc_feature    206250. .232406
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
ORIGIN

  Query Match             40.2%;  Score 160.6;  DB 12;  Length 232406;
  Best Local Similarity   77.0%;  Pred. No. 8.6e-44;
  Matches  207;  Conservative    1;  Mismatches   60;  Indels    1;  Gaps    1;

Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
              ||||||||||||||||||| |||||||||||  || |||||| |||||||| ||| | |
Db     149600 CCTGTAATCCCAGCACTTGGGGAGGCCAAGGCAGGTGGATCACTTGAGGTCAGGAGTTTG 149659

Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
              |||||| |||||||||| |||||||||| ||| | ||||||||||||||| |||||||
Db     149660 AGACCAGCCTGGCCAACACGGTGAAACCCCATCTCTCCTAAAAATACAAAAATTAGCTGG 149719

Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
              || |||||||  |||||| | |||||||||||||||||| | |||   | ||
Db     149720 GCCTGGTGGCATGCACCTGCAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGCTT 149779

Qy        264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
              || | |||   ||   ||| |    || | | | | || ||||||||||:|| ||||||
Db     149780 GAACCCAGGAGGACAAGAACTGCTTGAACCCTGGGCGA-CAGAGTGAGACTCCATCTCAA 149838

Qy        324 AAACAACAACAAAAAACAAAAAAACCATA 352
              ||| ||  | ||| || ||| |||||| |
Db     149839 AAAAAAACAAAAACAAAAAACAAACCAAA 149867


RESULT 5
AC005497
LOCUS       AC005497               97676 bp    DNA     linear   PRI 17-JUN-2000
DEFINITION  Homo sapiens chromosome 17, clone RP11-952N18, complete sequence.
ACCESSION   AC005497
VERSION     AC005497.9  GI:8570495
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 97676)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone RP11-952N18
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 97676)
  AUTHORS   Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
            Anderson,M., Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J.,
            Boutwell,C., Brown,A., Castle,A., Cerny,J., Colangelo,M.,
            Collins,S., Collymore,A., Cooke,P., Corliss,D., Depayre,E.,
            Devon,K., Dewar,K., Donelan,L., Ferreira,P., FitzHugh,W.,
            Forrest,C., Funke,R., Gage,D., Gardyna,S., Geraigery,K., Grant,G.,
            Hagos,B., Heaford,A., Herena,L., Horton,L., Howland,J.C.,
            Jacotot,L., Jones,C., Kann,L., Karatas,A., Lehoczky,J.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
            Nahf,R., Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P.,
            Pavlin,B., Peterson,K., Riley,R., Roberts,D., Roy,A.,
            Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C.,
            Subramanian,A., Tesfaye,S., Tichovolsky,N., Torruella-Miller,I.,
            Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,Y., Wyman,D.,
            Ye,W.J., Zhao,J. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-AUG-1998) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 97676)
```

```
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
          Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
          Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
          Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
          Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
          Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
          Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
          Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
          Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
          Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
          Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
          McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
          Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
          Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
          O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
          Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
          Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
          Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
          Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
          Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
          Young,G., Zainoun,J., Zimmer,A. and Zody,M.
TITLE     Direct Submission
JOURNAL   Submitted (17-JUN-2000) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT   On Jun 17, 2000 this sequence version replaced gi:7139014.
          All repeats were identified using RepeatMasker:
          Smit, A.F.A. & Green, P. (1996-1997)
          http://ftp.genome.washington.edu/RM/RepeatMasker.html

          Only 97676 base pairs from the middle of this clone are being
          submitted.  The remainder overlaps either accession AC015842 (WICGR
          project L449) or accession AC005304 (WICGR project L356).
          850 base pairs into the overlap with L449 [not submitted] is a VNTR
          that has been expanded by approximately 2.1 kilobases relative to
          project L449.
          -------------- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu
          -------------- Project Information
              Center project name: L370
              Center clone name: 952_N_18
          --------------.
FEATURES             Location/Qualifiers
     source          1. .97676
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-952N18"
                     /clone_lib="RPCI-11 Human Male BAC"
     repeat_region   complement(1. .38)
                     /rpt_family="MER5A"
     repeat_region   108. .217
                     /rpt_family="(GGAA)n"
     repeat_region   706. .851
                     /rpt_family="MIR"
     repeat_region   complement(1191. .1493)
                     /rpt_family="AluSx"
     repeat_region   complement(2304. .2474)
                     /rpt_family="MIR"
     repeat_region   2849. .3006
                     /rpt_family="MER82"
     repeat_region   3025. .3154
                     /rpt_family="Tigger2a"
     repeat_region   complement(3155. .3450)
                     /rpt_family="AluSp"
     repeat_region   3451. .3686
                     /rpt_family="Tigger2a"
     repeat_region   3687. .3795
                     /rpt_family="MER82"
     repeat_region   3796. .4096
                     /rpt_family="AluY"
     repeat_region   4097. .4155
                     /rpt_family="MER82"
     repeat_region   4156. .4526
                     /rpt_family="THE1C"
     repeat_region   4527. .4617
                     /rpt_family="MER82"
     repeat_region   5381. .5405
                     /rpt_family="(CA)n"
     repeat_region   5450. .6254
                     /rpt_family="L1M4"
     repeat_region   6267. .6324
                     /rpt_family="AT_rich"
     repeat_region   complement(6335. .6490)
                     /rpt_family="AluY"
     repeat_region   6659. .6829
                     /rpt_family="L1ME"
     repeat_region   6908. .7156
                     /rpt_family="L1ME"
     repeat_region   complement(7176. .7467)
                     /rpt_family="AluSx"
     repeat_region   8354. .8471
                     /rpt_family="MIR"
     repeat_region   8565. .8949
                     /rpt_family="L1MA7"
     repeat_region   9035. .9334
                     /rpt_family="AluSg"
     repeat_region   complement(9896. .10203)
                     /rpt_family="AluY"
     repeat_region   10650. .10960
                     /rpt_family="AluSq"
     repeat_region   complement(11015. .11035)
                     /rpt_family="MLT2FA"
     repeat_region   11036. .11055
                     /rpt_family="(CAAAA)n"
     repeat_region   complement(11056. .11493)
                     /rpt_family="MLT2FA"
     repeat_region   complement(12073. .12153)
                     /rpt_family="L2"
     repeat_region   complement(12154. .12318)
                     /rpt_family="MER5B"
     repeat_region   complement(12412. .12813)
                     /rpt_family="L2"
     repeat_region   12819. .13124
                     /rpt_family="AluSg"
     repeat_region   13289. .13579
                     /rpt_family="AluSg1"
     repeat_region   complement(13755. .14313)
                     /rpt_family="MLT2CB"
     repeat_region   14435. .14491
                     /rpt_family="AT_rich"
     repeat_region   14540. .14562
                     /rpt_family="AT_rich"
     repeat_region   15071. .15570
                     /rpt_family="L2"
     repeat_region   complement(15776. .16068)
                     /rpt_family="AluJo"
     repeat_region   16206. .16375
                     /rpt_family="FAM"
     repeat_region   complement(16490. .16803)
                     /rpt_family="AluY"
     repeat_region   17741. .17801
                     /rpt_family="AluJ/FRAM"
     repeat_region   complement(18123. .18563)
                     /rpt_family="L2"
     repeat_region   complement(18580. .18871)
                     /rpt_family="MLT1C"
     repeat_region   complement(18882. .19541)
                     /rpt_family="L2"
     repeat_region   19735. .19783
                     /rpt_family="L2"
     repeat_region   complement(19784. .20008)
```

```
                  /rpt_family="MER20"
     repeat_region   20009. .20057
                  /rpt_family="L2"
     repeat_region   complement(21315. .21633)
                  /rpt_family="MLT1G1"
     repeat_region   complement(21979. .22630)
                  /rpt_family="L1PA8"
     repeat_region   complement(23597. .23885)
                  /rpt_family="AluSc"
     repeat_region   24223. .24316
                  /rpt_family="MLT1E"
     repeat_region   24294. .24475
                  /rpt_family="MLT1E"
     repeat_region   24548. .24884
                  /rpt_family="MLT1E"
     repeat_region   25052. .25346
                  /rpt_family="AluSp"
     repeat_region   25415. .25448
                  /rpt_family="AT_rich"

  Query Match             40.0%;  Score 160;  DB 5;  Length 97676;
  Best Local Similarity   74.8%;  Pred. No. 9.3e-44;
  Matches  199;  Conservative    1;  Mismatches   66;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||||||||||||||| |||||||| |||  ||| ||||||||||||||  |||
Db       36133 AAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGGCAGATCACCTGAGGTCGGGAGG 36192

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || |||||| |||||||||||||| ||||| |||||| |||||||||| ||||| ||||
Db       36193 TCAAGACCAGCCTGGCCAACATGGAGAAACTCCGTCTCTACTAAAAATAGAAAAATTAGC 36252

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||||||||||  |||||| |||||||||||| |||| | |||   | ||
Db       36253 CGGGCATGGTGGCACATGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGGACTG 36312

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               || | |||   | | |    |||      | |  | |  ||||||| ||:|| |||
Db       36313 CTTGAACCCAGGAGGGGGAGGTTGCACTCCAGCCTGGTTAACAAGAGTGAAACTCCATCT 36372

Qy         321 CAAAAACAACAACAAAAAACAAAAAA 346
               || ||| || || ||||||.||||||
Db       36373 CACAAAAAAAAAAAAAAAAAAAAAA 36398


RESULT 6
AL445686/c
LOCUS       AL445686              107868 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-496D1 on chromosome 1 Contains
            the gene for a novel protein (FLJ42528), a ribosomal protein L26
            (RPL26) pseudogene, the 5' end of the SRRM1 gene for
            serine/arginine repetitive matrix 1 and a CpG island, complete
            sequence.
ACCESSION   AL445686
VERSION     AL445686.14  GI:12666298
KEYWORDS    HTG; FLJ42528; RPL26; SRRM1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 107868)
  AUTHORS   Coville,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 5, 2001 this sequence version replaced gi:12581050.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-496D1 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            ------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            -------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .107868
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-496D1"
                     /clone_lib="RPCI-11.2"
     misc_feature    100
                     /note="Clone_right_end: RP4-594I10"
     gene            join(9547. .9605,48860. .48948,54376. .54460,59042. .60672,
                     61043. .62761)
                     /locus_tag="RP11-496D1.1-001"
     mRNA            join(9547. .9605,48860. .48948,54376. .54460,59042. .60672,
                     61043. .62761)
                     /locus_tag="RP11-496D1.1-001"
                     /product="novel protein"
                     /note="match: cDNAs: Em:AK124519.1"
     gene            17036. .17402
                     /locus_tag="RP11-496D1.2-001"
                     /pseudo
     CDS             17036. .17402
                     /locus_tag="RP11-496D1.2-001"
                     /note="match: proteins: Sw:P12749 Sw:Q02877 Sw:Q95WA0
                     Sw:Q9UNX3 Tr:BAC21653 Tr:BAC56365"
                     /pseudo
                     /codon_start=1
                     /product="ribosomal protein L26 (RPL26) pseudogene"
     misc_feature    38568
                     /locus_tag="RP11-496D1.1-001"
                     /note="Clone_left_end: RP11-373M8"
     gene            join(42900. .43060,48860. .48953,54376. .54460,59042. .60466,
                     62295. .62764)
                     /locus_tag="RP11-496D1.1-002"
     mRNA            join(42900. .43060,48860. .48953,54376. .54460,59042. .60466,
                     62295. .62764)
                     /locus_tag="RP11-496D1.1-002"
                     /product="novel protein"
                     /note="match: cDNAs: Em:BC050474.1"
     CDS             join(48867. .48948,54376. .54460,59042. .59183)
                     /locus_tag="RP11-496D1.1-001"
                     /standard_name="OTTHUMP00000003372"
                     /codon_start=1
                     /product="novel protein"
                     /protein_id="CAI14681.1"
                     /db_xref="GI:55959942"
                     /db_xref="UniProtKB/TrEMBL:Q5T1S8"
                     /translation="MTTATPLGDTTFFSLNMTTRGEDFLYKSSGAIVAAVVVVVIIIF
                     TVVLILLKMYNRKMRTRRELEPKGPKPTAPSAVGPNSNGSQHPATVTFSPVDVQVETR
                     "
     polyA_signal    62744. .62749
                     /locus_tag="RP11-496D1.1-002"
```

```
polyA_site      62764
                /locus_tag="RP11-496D1.1-002"
gene            join(85152. .85213,86455. .86535,89080. .89313)
                /gene="SRRM1"
                /locus_tag="RP11-373M8.1-005"
mRNA            join(85152. .85213,86455. .86535,89080. .89313)
                /gene="SRRM1"
                /locus_tag="RP11-373M8.1-005"
                /product="serine\/arginine repetitive matrix 1"
                /note="match: ESTs: Em:AI123635.1"
gene            join(86502. .86535,99420. .99509,100103. .100225,
                102295. .102465,103407. .103468)
                /gene="SRRM1"
                /locus_tag="RP11-373M8.1-006"
mRNA            join(86502. .86535,99420. .99509,100103. .100225,
                102295. .102465,103407. .103468)
                /gene="SRRM1"
                /locus_tag="RP11-373M8.1-006"
                /product="serine\/arginine repetitive matrix 1"
                /note="match: ESTs: Em:BQ020684.1"
gene            join(96539. .96783,99420. .99509,100103. .100225,
                102295. .102465,103407. .103522,104845. .105048,
                105870. .106064,106349. .106468,AL445648.18:2423. .2697,
                AL445648.18:8287. .8367,AL445648.18:8878. .8964,
                AL445648.18:10228. .10372,AL445648.18:14383. .14493,
                AL445648.18:16691. .17155,AL445648.18:17688. .17883,
                AL445648.18:18954. .19163,AL445648.18:19750. .20835)
                /gene="SRRM1"
                /locus_tag="RP11-373M8.1-001"
mRNA            join(96539. .96783,99420. .99509,100103. .100225,
                102295. .102465,103407. .103522,104845. .105048,
                105870. .106064,106349. .106468,AL445648.18:2423. .2697,
                AL445648.18:8287. .8367,AL445648.18:8878. .8964,
                AL445648.18:10228. .10372,AL445648.18:14383. .14493,
                AL445648.18:16691. .17155,AL445648.18:17688. .17883,
                AL445648.18:18954. .19163,AL445648.18:19750. .20835)
                /gene="SRRM1"
                /locus_tag="RP11-373M8.1-001"
                /product="serine\/arginine repetitive matrix 1"
                /note="match: cDNAs: Em:AF419855.1 Em:BC017315.1
                Em:BC036187.1"
gene            join(96708. .96783,99420. .99509,100103. .100225,
                100515. .101024)
                /gene="SRRM1"
                /locus_tag="RP11-373M8.1-007"
mRNA            join(96708. .96783,99420. .99509,100103. .100225,
                100515. .101024)
                /gene="SRRM1"
                /locus_tag="RP11-373M8.1-007"
                /product="serine\/arginine repetitive matrix 1"
                /note="match: ESTs: Em:AL693732.1 Em:AL704047.1
                Em:AW964246.1"
gene            join(96709. .96783,99420. .99509,100103. .100225,
                102295. .102391,103470. .103522,104845. .105048,
                105870. .105978)
                /gene="SRRM1"
                /locus_tag="RP11-373M8.1-008"
mRNA            join(96709. .96783,99420. .99509,100103. .100225,
                102295. .102391,103470. .103522,104845. .105048,
                105870. .105978)
                /gene="SRRM1"
                /locus_tag="RP11-373M8.1-008"
                /product="serine\/arginine repetitive matrix 1"
                /note="match: ESTs: Em:AU142070.1"
gene            join(96711. .96783,99420. .99509,100103. .100225,
                102295. .102465,103407. .103522,104935. .105048,
                105870. .105950)
                /gene="SRRM1"
                /locus_tag="RP11-373M8.1-009"
mRNA            join(96711. .96783,99420. .99509,100103. .100225,
                102295. .102465,103407. .103522,104935. .105048,
                105870. .105950)
                /gene="SRRM1"
                /locus_tag="RP11-373M8.1-009"
                /product="serine\/arginine repetitive matrix 1"
                /note="match: ESTs: Em:BG700853.1"
gene            join(96729. .96783,99420. .99509,100103. .102465,
                103407. .103522)
                /gene="SRRM1"
                /locus_tag="RP11-373M8.1-004"
mRNA            join(96729. .96783,99420. .99509,100103. .102465,
                103407. .103522)
                /gene="SRRM1"
                /locus_tag="RP11-373M8.1-004"
                /product="serine\/arginine repetitive matrix 1"
                /note="match: ESTs: Em:CD656593.1 Em:CD706723.1
                match: cDNAs: Em:AK032599.1"
gene            join(96735. .96783,99420. .99509,100103. .100225,
                102295. .102465,103407. .103522,104845. .105048,
                105870. .106064,106364. .106468,AL445648.18:2423. .2697,
                AL445648.18:8287. .8367,AL445648.18:8878. .8964,
                AL445648.18:10228. .10372,AL445648.18:10751. .10792,
                AL445648.18:14383. .14493,AL445648.18:16691. .17155,
                AL445648.18:17688. .17883,AL445648.18:18954. .19163,
                AL445648.18:19750. .20830)
                /gene="SRRM1"
                /locus_tag="RP11-373M8.1-002"
mRNA            join(96735. .96783,99420. .99509,100103. .100225,
                102295. .102465,103407. .103522,104845. .105048,
                105870. .106064,106364. .106468,AL445648.18:2423. .2697,
                AL445648.18:8287. .8367,AL445648.18:8878. .8964,
```

```
  Query Match             40.0%;  Score 159.8;  DB 5;  Length 107868;
  Best Local Similarity   71.8%;  Pred. No. 1.1e-43;
  Matches  209;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||||| |||||||  |||  || || ||||||||||||||  ||
Db      22711 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAGGTGGGTCACCTGAGGTCAGAAGT 22652

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||| |||||||||||||||||||||||||||| ||||||||||||||||  ||
Db      22651 TTGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATCAGT 22592

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||||||||||  | |||||| || ||||||||| |||| | ||| | | ||
Db      22591 CGGGCATGGTGGCAGGCGCCTGTAATCTCAGCTACTCGGGAGGCTGAGGTGGGAGAATTG 22532

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              || | |||||   || |  | |  ||||| ||   |  |  ||  |||       ||
Db      22531 CTTGACCCCAGAGGTTGCAGTGAGCTGAGATCGCACCATTGGGCGACAGAGCAAGTCTCC 22472

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGG 371
              ||||| ||||| | |||||||||| ||  |||  |     |  | ||| ||
Db      22471 AAAAAAAAACAAAACAAAACAAAACAAATATATTAGGAATTAAAGCTGTGG 22421
```

```
RESULT 7
AC016542
LOCUS       AC016542              171480 bp    DNA     linear   PRI 09-APR-2002
DEFINITION  Homo sapiens chromosome 10 clone RP11-354E23, complete sequence.
ACCESSION   AC016542
VERSION     AC016542.11  GI:19744959
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 171480)
  AUTHORS   Smith,D.R.
  TITLE     Genome Therapeutics Corporation Sequencing Center: Human Genome
            Sequence Data
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 171480)
  AUTHORS   Smith,D.R.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-1999) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   3  (bases 1 to 171480)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-2001) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   4  (bases 1 to 171480)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2002) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   5  (bases 1 to 171480)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAR-2002) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   6  (bases 1 to 171480)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-APR-2002) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
COMMENT     On Mar 27, 2002 this sequence version replaced gi:16930889.
            -------------- Genome Center
                Center: Genome Therapeutics Corporation
                Center code: GTC
                Web site: http://www.genomecorp.com/
                Contact: gtc-seqcenter@genomecorp.com
            -------------- Project Information
                Center project name: hg069

                IMPORTANT: This sequence is the entire
                insert of clone RP11-354E23.
                The true right end of clone RP11-522H2 is at
                2000 in this sequence.
            -------------- Summary Statistics
                Sequencing vector: N/A
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 990315
            --------------.
FEATURES             Location/Qualifiers
    source           1. .171480
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-354E23"
                     /clone_lib="RPCI-11"
ORIGIN

  Query Match             40.0%;  Score 159.8;  DB 5;  Length 171480;
  Best Local Similarity   71.8%;  Pred. No. 1.4e-43;
  Matches  222;  Conservative    1;  Mismatches   83;  Indels    3;  Gaps    1;

Qy          84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
               |||||||||||||||| |||||||| |||||||||||||||||||||||| ||| |||
Db      116796 CATGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCGGATCACCTGAGGTCAGGAGTTCG 116855

Qy         144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
               |||||| ||||| |||||||||||||||||||| |||||||||| ||||| |||||||
Db      116856 AGACCAGCCTGGGCAACATGGTGAAACCCCGTCTCTACTAAAAATATAAAAATTAGCTGG 116915

Qy         204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
               ||||||||||    |||||| |||||||||| | |||| | |||   | ||
Db      116916 GCATGGTGGCATGTGCCTGTAATCCCAGCTACCCGGGAGGCTGAGGCAGGAGAATAGCTT 116975

Qy         264 GATCGCAGAGTGAGCCGAAATCACAGATCACAG---AGTGAGCAGAGTGAGACKCCGTCT 320
               || | |||   | |      |     ||   ||   ||    ||||||||||:||||||
Db      116976 GAACCCAGGAGGCGGATTGGTTCTGCATTCTAGCCTAGGTGACAGAGTGAGACTCCGTCT 117035

Qy         321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGAC 380
               |||||| || || ||||||  |||||   |  |   || |   | |   | ||   |||

Db      117036 CAAAAAAAAAAAAAAAAAAAGAAAAAGAAAGGACCAATAGGAGAACAAAGTTTAAAAGAA 117095

Qy         381 TATTGCAGG 389
               | ||| |||
Db      117096 TGTTGAAGG 117104


RESULT 8
HS408N23
LOCUS       HS408N23               97916 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP3-408N23 on chromosome 22q13,
            complete sequence.
ACCESSION   Z98048
VERSION     Z98048.1  GI:2582746
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 97916)
  AUTHORS   Hunt,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 2, 1997 this sequence version replaced gi:2462400.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 22, constructed by the Sanger Centre Chromosome 22
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr22
            RP3-408N23 is from the library RPCI-3 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .97916
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704N23408"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /map="q13"
                     /clone="RP3-408N23"
                     /clone_lib="RPCI-3"
     gene            complement(2630. .4351)
                     /locus_tag="RP3-408N23.3-003"
                     /pseudo
     CDS             complement(2630. .4351)
                     /locus_tag="RP3-408N23.3-003"
                     /note="match: proteins: Tr:O76206"
                     /pseudo
                     /codon_start=1
     gene            join(complement(7439. .34299),
                     complement(AL049764.4:12691. .22186))
```

```
                     /locus_tag="RP3-362J20.1-001"
     mRNA            join(complement(34109. .34299),complement(13937. .13997),
                     complement(9428. .9494),complement(7439. .7590),
                     complement(AL049764.4:22070. .22186),
                     complement(AL049764.4:20289. .20434),
                     complement(AL049764.4:20101. .20196),
                     complement(AL049764.4:16998. .17080),
                     complement(AL049764.4:12691. .14042))
                     /locus_tag="RP3-362J20.1-001"
     CDS             join(complement(34109. .34162),complement(13937. .13997),
                     complement(9428. .9494),complement(7439. .7590),
                     complement(AL049764.4:22070. .22186),
                     complement(AL049764.4:20289. .20434),
                     complement(AL049764.4:20101. .20196),
                     complement(AL049764.4:16998. .17080),
                     complement(AL049764.4:13895. .14042))
                     /locus_tag="RP3-362J20.1-001"
                     /standard_name="OTTHUMP00000028507"
                     /codon_start=1
                     /protein_id="CAI20491.1"
                     /db_xref="GI:56208218"
                     /db_xref="InterPro:IPR001993"
                     /db_xref="InterPro:IPR002030"
                     /db_xref="InterPro:IPR002067"
                     /db_xref="InterPro:IPR002113"
                     /translation="MASVLSYESLVHAVAGAVGSVTAMTVFFPLDTARLRLQVDEKRK
                     SKTTHMVLLEIIKEEGLLAPYRGWFPVISSLCCSNFVYFYTFNSLKALWVKGQHSTTG
                     KDLVVGFVAGVVNVLLTTPLWVVNTRLKLQGAKFRNEDIVPTNYKGIIDAFHQIIRDE
                     GISALWNGTFPSLLLVFNPAIQFMFYEGLKRQLLKKRMKLSSLDVFIIGAVAKAIATT
                     VTYPLQTVQSILRFGRHRLNPENRTLGSLRNILYLLHQRVRRFGIMGLYKGLEAKLLQ
                     TVLTAALMFLVYEKLTAATFTVMGLKRAHQH"
     gene            complement(39513. .71758)
                     /locus_tag="RP3-408N23.1-001"
     mRNA            complement(join(39513. .41580,42010. .42143,44540. .44588,
                     45754. .45870,47483. .47585,50474. .50584,50717. .50801,
                     55540. .55606,59753. .59823,63208. .63283,65737. .65794,
                     71345. .71758))
                     /locus_tag="RP3-408N23.1-001"
     CDS             complement(join(41452. .41580,42010. .42143,44540. .44588,
                     45754. .45870,47483. .47585,50474. .50584,50717. .50801,
                     55540. .55606,59753. .59823,63208. .63283,65737. .65794,
                     71345. .71454))
                     /locus_tag="RP3-408N23.1-001"
                     /standard_name="OTTHUMP00000028873"
                     /codon_start=1
                     /protein_id="CAB10844.1"
                     /db_xref="GI:2598461"
                     /db_xref="GOA:P50502"
                     /db_xref="HGNC:11343"
                     /db_xref="InterPro:IPR001440"
                     /db_xref="InterPro:IPR006636"
                     /db_xref="UniProtKB/Swiss-Prot:P50502"
                     /translation="MDPRKVNELRAFVKMCKQDPSVLHTEEMRFLREWVESMGGKVPP
                     ATQKAKSEENTKEEKPDSKKVEEDLKADEPSSEESDLEIDKEGVIEPDTDAPQEMGDE
                     NAEITEEMMDQANDKKVAAIEALNDGELQKAIDLFTDAIKLNPRLAILYAKRASVFVK
                     LQKPNAAIRDCDRAIEINPDSAQPYKWRGKAHRLLGHWEEAAHDLALACKLDYDEDAS
                     AMLKEVQPRAQKIAEHRRKYERKREEREIKERIERVKKAREEHERAQREEEARRQSGA
                     QYGSFPGGFPGGMPGNFPGGMPGMGGGMPGMAGMPGLNEILSDPEVLAAMQDPEVMVA
                     FQDVAQNPANMSKYQSNPKVMNLISKLSAKFGGQA"
     gene            join(72012. .72159,83913. .84029,96684. .97091,
                     AL035450.1:3415. .3617,AL035450.1:24689. .24751,
                     AL035450.1:26224. .26337,AL035450.1:31341. .31426,
                     AL035450.1:39435. .39615,AL035450.1:41464. .41584,
                     AL035450.1:43371. .46534,AL080242.11:105. .3488)
                     /locus_tag="RP5-1057D18.1-001"
     mRNA            join(72012. .72159,83913. .84029,96684. .97091,
                     AL035450.1:3415. .3617,AL035450.1:24689. .24751,
                     AL035450.1:26224. .26337,AL035450.1:31341. .31426,
                     AL035450.1:39435. .39615,AL035450.1:41464. .41584,
                     AL035450.1:43371. .46534,AL080242.11:105. .3488)
                     /locus_tag="RP5-1057D18.1-001"
     CDS             join(72096. .72159,83913. .84029,96684. .97091,
                     AL035450.1:3415. .3617,AL035450.1:24689. .24751,
                     AL035450.1:26224. .26337,AL035450.1:31341. .31426,
                     AL035450.1:39435. .39615,AL035450.1:41464. .41584,
                     AL035450.1:43371. .46534)
                     /locus_tag="RP5-1057D18.1-001"
                     /standard_name="OTTHUMP00000028904"
                     /codon_start=1
                     /protein_id="CAI20492.1"
                     /db_xref="GI:56208219"
                     /db_xref="GOA:Q9NQH7"
                     /db_xref="InterPro:IPR000994"
                     /db_xref="InterPro:IPR007865"
                     /db_xref="UniProtKB/TrEMBL:Q9NQH7"
                     /translation="MPWLLSAPKLVPAVANVRGLSGCMLCSQRRYSLQPVPERRIPNR
                     YLGQPSPFTHPHLLRPGEVTPGLSQVEYALRRHKLMSLIQKEAQGQSGTDQTVVVLSN
                     PTYYMSNDIPYTFHQDNNFLYLCGFQEPDSILVLQSLPGKQLPSHKAILFVPRRDPSR
                     ELWDGPRSGTDGAIALTGVDEAYTLEEFQHLLPKMKAETNMVWYDWMRPSHAQLHSDY
                     MQPLTEAKAKSKNKVRGVQQLIQRLRLIKSPAEIERMQIAGKLTSQAFIETMFTSKAP
                     VEEAFLYAKFEFECRARGADILAYPPVVAGGNRSNTLHYVKNNQLIKDGEMVLLDGGC
                     ESSCYVSDITRTWPVNGRFTAPQAELYEAVLEIQRDCLALCFPGTSLENIYSMMLTLI
                     GQKLKDLGIMKNIKENNAFKAARKYCPHHVGHYLGMDVHDTPDMPRSLPLQPGMVITI
                     EPGIYIPEDDKDAPEKFRGLGVRIEDDVVVTQDSPLILSADCPKEMNDIEQICSQAS"
     gene            complement(74463. .77040)
                     /locus_tag="RP3-408N23.2-002"
     mRNA            complement(74463. .77040)
                     /locus_tag="RP3-408N23.2-002"
     CDS             complement(75979. .76908)
                     /locus_tag="RP3-408N23.2-002"
                     /standard_name="OTTHUMP00000028874"
                     /note="match: proteins: Tr:O54946"
                     /codon_start=1
                     /protein_id="CAI20493.1"
                     /db_xref="GI:56208220"
                     /db_xref="InterPro:IPR001623"
                     /translation="MVDYYEVLGLQRYASPEDIKKAYHKVALKWHPDKNPENKEEAER
                     KFKEVAEAYEVLSNDEKRDIYDKYGTEGLNGGGSHFDDECEYGFTFHKPDDVFKEIFH
                     ERDPFSFHFFEDSLEDLLNRPGSSYGNRNRDAGYFFSTASEYPIFEKFSSYDTGYTSQ
                     GSLGHEGLTSFSSLAFDNSGMDNYISVTTSDKIVNGRNINTKKIIESDQEREAEDNGE
                     LTFFLVNSVANEEGFAKECSWRTQSFNNYSPNSHSSKHVSQYTFVDNDEGGISWVTSN
                     RDPPIFSAGVKEGGKRKKKKRKEVQKKSTKRNC"
ORIGIN

  Query Match             39.9%;  Score 159.4;  DB 5;  Length 97916;
  Best Local Similarity   75.7%;  Pred. No. 1.5e-43;
  Matches  196;  Conservative    1;  Mismatches   62;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||  |||||||  ||||||||||||||||| |||
Db      11289 ACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGTGGATCACCTGAGGTCAGGAGT 11348

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              |||||| || |||||||||| ||| |||||||| ||| ||||||||||||||||| ||||
Db      11349 TCGAGATCAGCCTGGCCAACGTGGCGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 11408

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||||||||||||||  |||||| |||||||||||| |||| | ||||  | ||
Db      11409 TGGGCATGGTGGCACATGCCTGTAATCCCAGCTACTCGGGAGGCTGAGACAGGAGAATCA 11468

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              | ||   |||   |||        ||| |  | |  |     |   || |||||:| ||||
Db      11469 CGTGAATCCAGGCCGAGATTGTGCCACTGCACTCCAACCTGGGTAAGAGAGACTCTGTCT 11528

Qy        321 CAAAAACAACAACAAAAAA 339
              |||||| || || |||| |
Db      11529 CAAAAAAAAAAAAAAAGA 11547


RESULT 9
AC090344/c
LOCUS       AC090344              195616 bp    DNA     linear   HTG 07-JUL-2001
DEFINITION  Homo sapiens chromosome 11 clone RP11-804A23 map 11, WORKING DRAFT
            SEQUENCE, 10 unordered pieces.
ACCESSION   AC090344
VERSION     AC090344.3  GI:14626339
```

```
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 195616)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 11, clone RP11-804A23
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 195616)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
            Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
            Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos;B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
            Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
            Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
            Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
            Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
            Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jul 7, 2001 this sequence version replaced gi:13357354.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L12622
                Center clone name: 804_A_23
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 191318 bases at least Q40
                Consensus quality: 193221 bases at least Q30
                Consensus quality: 194120 bases at least Q20
                Insert size: 194716; sum-of-contigs
                Quality coverage: 9.4 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 10 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     9269: contig of 9269 bp in length
            *     9270     9369: gap of 100 bp
            *     9370    10358: contig of 989 bp in length
            *    10359    10458: gap of 100 bp
            *    10459    13550: contig of 3092 bp in length
            *    13551    13650: gap of 100 bp
            *    13651    26222: contig of 12572 bp in length
            *    26223    26322: gap of 100 bp
            *    26323    86422: contig of 60100 bp in length
            *    86423    86522: gap of 100 bp
            *    86523   100998: contig of 14476 bp in length
            *   100999   101098: gap of 100 bp
            *   101099   120502: contig of 19404 bp in length
            *   120503   120602: gap of 100 bp
            *   120603   146265: contig of 25663 bp in length
            *   146266   146365: gap of 100 bp
            *   146366   180564: contig of 34199 bp in length
            *   180565   180664: gap of 100 bp
            *   180665   195616: contig of 14952 bp in length.
FEATURES             Location/Qualifiers
     source          1. .195616
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11"
                     /clone="RP11-804A23"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .9269
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     gap             9270. .9369
                     /estimated_length=100
     misc_feature    9370. .10358
                     /note="assembly_fragment"
     gap             10359. .10458
                     /estimated_length=100
     misc_feature    10459. .13550
                     /note="assembly_fragment"
     gap             13551. .13650
                     /estimated_length=100
     misc_feature    13651. .26222
                     /note="assembly_fragment"
     gap             26223. .26322
                     /estimated_length=100
     misc_feature    26323. .86422
                     /note="assembly_fragment"
     gap             86423. .86522
                     /estimated_length=100
     misc_feature    86523. .100998
                     /note="assembly_fragment"
     gap             100999. .101098
                     /estimated_length=100
     misc_feature    101099. .120502
                     /note="assembly_fragment"
     gap             120503. .120602
                     /estimated_length=100
     misc_feature    120603. .146265
                     /note="assembly_fragment"
     gap             146266. .146365
                     /estimated_length=100
     misc_feature    146366. .180564
                     /note="assembly_fragment"
     gap             180565. .180664
                     /estimated_length=100
     misc_feature    180665. .195616
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             39.7%;  Score 158.8;  DB 12;  Length 195616;
  Best Local Similarity   74.0%;  Pred. No. 3.5e-43;
  Matches  216;  Conservative    1;  Mismatches   68;  Indels    7;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||||| |||||||| ||| |||||||||| |||||||| |||
Db      91667 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGGCGGATCACTTGAGGTCAGGAGT 91608

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
```

```
            ||||||||| |||||||||||| |||||||||| || |||||||||||||||| ||||
Db      91607 TCGAGACCAGCCTGGCCAACATGATGAAACCCCGCCTCTACTAAAAATACAAAAATTAGC 91548

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-------GCCGGAGATTGC 253
              ||| ||||||||     |||||| |||||||||||| |||       || ||||| |
Db      91547 CAGGCGTGGTGGCAGTTGCCTGTAATCCCAGCTACTCGGGAGACTGAGGCAGGAGAATTG 91488

Qy        254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGAC 313
              | |||  ||   ||||||   || | || |  |||||       |   ||||||||||
Db      91487 CTTAAGCCCAGGAGGCAGAGGTTGCAGTAAGCTTAGATCGTGCCTGGGCAAGAGTGAGAC 91428

Qy        314 KCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCAT 365
              :|| |||||| || || || |||||  || |||||   ||    | ||| ||
Db      91427 TCCATCTCAAGAAAAAAAAAAAAAAGAAAGAAAACAGAAACCTGTAGTCTAT 91376


RESULT 10
AC090345/c
LOCUS       AC090345               197156 bp    DNA     linear   HTG 11-JUL-2001
DEFINITION  Homo sapiens chromosome 11 clone RP11-804B24 map 11, WORKING DRAFT
            SEQUENCE, 10 unordered pieces.
ACCESSION   AC090345
VERSION     AC090345.3  GI:14670111
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 197156)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 11, clone RP11-804B24
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 197156)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
            Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
            Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
            Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
            Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
            Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
            Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
            Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jul 11, 2001 this sequence version replaced gi:13357355.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L12623
                Center clone name: 804_B_24
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 193219 bases at least Q40
                Consensus quality: 194831 bases at least Q30
                Consensus quality: 195508 bases at least Q20
                Insert size: 196256; sum-of-contigs
                Quality coverage: 10.0 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 10 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    11369: contig of 11369 bp in length
            *    11370    11469: gap of 100 bp
            *    11470    14984: contig of 3515 bp in length
            *    14985    15084: gap of 100 bp
            *    15085    21800: contig of 6716 bp in length
            *    21801    21900: gap of 100 bp
            *    21901    61396: contig of 39496 bp in length
            *    61397    61496: gap of 100 bp
            *    61497    83042: contig of 21546 bp in length
            *    83043    83142: gap of 100 bp
            *    83143    97779: contig of 14637 bp in length
            *    97780    97879: gap of 100 bp
            *    97880   124355: contig of 26476 bp in length
            *   124356   124455: gap of 100 bp
            *   124456   150156: contig of 25701 bp in length
            *   150157   150256: gap of 100 bp
            *   150257   182157: contig of 31901 bp in length
            *   182158   182257: gap of 100 bp
            *   182258   197156: contig of 14899 bp in length.
FEATURES             Location/Qualifiers
     source          1. .197156
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11"
                     /clone="RP11-804B24"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .11369
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     gap             11370. .11469
                     /estimated_length=100
     misc_feature    11470. .14984
                     /note="assembly_fragment"
     gap             14985. .15084
                     /estimated_length=100
     misc_feature    15085. .21800
                     /note="assembly_fragment"
     gap             21801. .21900
                     /estimated_length=100
     misc_feature    21901. .61396
                     /note="assembly_fragment"
     gap             61397. .61496
                     /estimated_length=100
     misc_feature    61497. .83042
                     /note="assembly_fragment"
     gap             83043. .83142
                     /estimated_length=100
     misc_feature    83143. .97779
                     /note="assembly_fragment"
     gap             97780. .97879
                     /estimated_length=100
     misc_feature    97880. .124355
                     /note="assembly_fragment"
     gap             124356. .124455
                     /estimated_length=100
```

```
     misc_feature    124456. .150156
                     /note="assembly_fragment"
     gap             150157. .150256
                     /estimated_length=100
     misc_feature    150257. .182157
                     /note="assembly_fragment"
     gap             182158. .182257
                     /estimated_length=100
     misc_feature    182258. .197156
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             39.7%;  Score 158.8;  DB 12;  Length 197156;
  Best Local Similarity   74.0%;  Pred. No. 3.5e-43;
  Matches  216;  Conservative    1;  Mismatches   68;  Indels    7;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              ||||||||||||||||||||| |||||||| |||  |||||||||| |||||||| |||
Db      88380 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGGCGGATCACTTGAGGTCAGGAGT 88321

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              ||||||||| |||||||||||| |||||||||| || ||||||||||||||||| ||||
Db      88320 TCGAGACCAGCCTGGCCAACATGATGAAACCCCGCCTCTACTAAAAATACAAAAATTAGCC 88261

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-------GCCGGAGATTGC 253
                |||  ||||||||     |||||| |||||||||||| |||        || ||||| |
Db      88260 CAGGCGTGGTGGCAGTTGCCTGTAATCCCAGCTACTCGGGAGACTGAGGCAGGAGAATTG 88201

Qy        254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGAC 313
                |.|||  ||   ||||||   || | || |  |||||        |   ||||||||||
Db      88200 CTTAAGCCCAGGAGGCAGAGGTTGCAGTAAGCTTAGATCGTGCCTGGGCAAGAGTGAGAC 88141

Qy        314 KCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCAT 365
              :|| |||||| || || || ||||| || |||||   ||    | ||| ||
Db      88140 TCCATCTCAAGAAAAAAAAAAAAAAGAAAGAAAACAGAAACCTGTAGTCTAT 88089


RESULT 11
AP000777
LOCUS       AP000777               197856 bp    DNA     linear   PRI 01-FEB-2003
DEFINITION  Homo sapiens genomic DNA, chromosome 11 clone:RP11-804A23, complete
            sequence.
ACCESSION   AP000777
VERSION     AP000777.4  GI:28189489
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA
  JOURNAL   Published Only in Database (1999)
REFERENCE   2  (bases 1 to 197856)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-NOV-1999) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     On Jan 31, 2003 this sequence version replaced gi:23821512.
FEATURES             Location/Qualifiers
     source          1. .197856
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q"
                     /clone="RP11-804A23"
ORIGIN

  Query Match             39.7%;  Score 158.8;  DB 5;  Length 197856;
  Best Local Similarity   74.0%;  Pred. No. 3.5e-43;
  Matches  216;  Conservative    1;  Mismatches   68;  Indels    7;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              ||||||||||||||||||||| |||||||| |||  |||||||||| |||||||| |||
Db      52629 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGGCGGATCACTTGAGGTCAGGAGT 52688

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              ||||||||| |||||||||||| |||||||||| || ||||||||||||||||| ||||
Db      52689 TCGAGACCAGCCTGGCCAACATGATGAAACCCCGCCTCTACTAAAAATACAAAAATTAGCC 52748

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-------GCCGGAGATTGC 253
                |||  ||||||||     |||||| |||||||||||| |||        || ||||| |
Db      52749 CAGGCGTGGTGGCAGTTGCCTGTAATCCCAGCTACTCGGGAGACTGAGGCAGGAGAATTG 52808

Qy        254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGAC 313
                | |||  ||   ||||||   || | || |  |||||        |   ||||||||||
Db      52809 CTTAAGCCCAGGAGGCAGAGGTTGCAGTAAGCTTAGATCGTGCCTGGGCAAGAGTGAGAC 52868

Qy        314 KCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCAT 365
              :|| |||||| || || || ||||| || |||||   ||    | ||| ||
Db      52869 TCCATCTCAAGAAAAAAAAAAAAAAGAAAGAAAACAGAAACCTGTAGTCTAT 52920


RESULT 12
AC087662/c
LOCUS       AC087662               68990 bp    DNA     linear   HTG 15-JAN-2001
DEFINITION  Homo sapiens chromosome 8 clone RP5-1155K23 map 8, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC087662
VERSION     AC087662.1  GI:12229437
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 68990)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 8, clone RP5-1155K23
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 68990)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
            Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
            Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
            Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
            Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
            Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
            Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
            Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
```

```
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-------------- Genome Center
    Center: Whitehead Institute/ MIT Center for Genome Research
    Center code: WIBR
    Web site: http://www-seq.wi.mit.edu
    Contact: sequence_submissions@genome.wi.mit.edu
-------------- Project Information
    Center project name: L12144
    Center clone name: 1155_K_23
--------------.
* NOTE: This record contains 84 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*        1      709: contig of 709 bp in length
*      710      809: gap of 100 bp
*      810     1546: contig of 737 bp in length
*     1547     1646: gap of 100 bp
*     1647     2377: contig of 731 bp in length
*     2378     2477: gap of 100 bp
*     2478     3197: contig of 720 bp in length
*     3198     3297: gap of 100 bp
*     3298     4024: contig of 727 bp in length
*     4025     4124: gap of 100 bp
*     4125     4842: contig of 718 bp in length
*     4843     4942: gap of 100 bp
*     4943     5652: contig of 710 bp in length
*     5653     5752: gap of 100 bp
*     5753     6491: contig of 739 bp in length
*     6492     6591: gap of 100 bp
*     6592     7317: contig of 726 bp in length
*     7318     7417: gap of 100 bp
*     7418     8143: contig of 726 bp in length
*     8144     8243: gap of 100 bp
*     8244     8968: contig of 725 bp in length
*     8969     9068: gap of 100 bp
*     9069     9740: contig of 672 bp in length
*     9741     9840: gap of 100 bp
*     9841    10580: contig of 740 bp in length
*    10581    10680: gap of 100 bp
*    10681    11406: contig of 726 bp in length
*    11407    11506: gap of 100 bp
*    11507    12216: contig of 710 bp in length
*    12217    12316: gap of 100 bp
*    12317    13047: contig of 731 bp in length
*    13048    13147: gap of 100 bp
*    13148    13878: contig of 731 bp in length
*    13879    13978: gap of 100 bp
*    13979    14711: contig of 733 bp in length
*    14712    14811: gap of 100 bp
*    14812    15525: contig of 714 bp in length
*    15526    15625: gap of 100 bp
*    15626    16323: contig of 698 bp in length
*    16324    16423: gap of 100 bp
*    16424    17156: contig of 733 bp in length
*    17157    17256: gap of 100 bp
*    17257    17969: contig of 713 bp in length
*    17970    18069: gap of 100 bp
*    18070    18795: contig of 726 bp in length
*    18796    18895: gap of 100 bp
*    18896    19627: contig of 732 bp in length
*    19628    19727: gap of 100 bp
*    19728    20447: contig of 720 bp in length
*    20448    20547: gap of 100 bp
*    20548    21267: contig of 720 bp in length
*    21268    21367: gap of 100 bp
*    21368    22085: contig of 718 bp in length
*    22086    22185: gap of 100 bp
*    22186    22883: contig of 698 bp in length
*    22884    22983: gap of 100 bp
*    22984    23711: contig of 728 bp in length
*    23712    23811: gap of 100 bp
*    23812    24522: contig of 711 bp in length
*    24523    24622: gap of 100 bp
*    24623    25334: contig of 712 bp in length
*    25335    25434: gap of 100 bp
*    25435    26157: contig of 723 bp in length
*    26158    26257: gap of 100 bp
*    26258    26973: contig of 716 bp in length
*    26974    27073: gap of 100 bp
*    27074    27797: contig of 724 bp in length
*    27798    27897: gap of 100 bp
*    27898    28618: contig of 721 bp in length
*    28619    28718: gap of 100 bp
*    28719    29466: contig of 748 bp in length
*    29467    29566: gap of 100 bp
*    29567    30295: contig of 729 bp in length
*    30296    30395: gap of 100 bp
*    30396    31112: contig of 717 bp in length
*    31113    31212: gap of 100 bp
*    31213    31934: contig of 722 bp in length
*    31935    32034: gap of 100 bp
*    32035    32750: contig of 716 bp in length
*    32751    32850: gap of 100 bp
*    32851    33571: contig of 721 bp in length
*    33572    33671: gap of 100 bp
*    33672    34404: contig of 733 bp in length
*    34405    34504: gap of 100 bp
*    34505    35212: contig of 708 bp in length
*    35213    35312: gap of 100 bp
*    35313    36048: contig of 736 bp in length
*    36049    36148: gap of 100 bp
*    36149    36859: contig of 711 bp in length
*    36860    36959: gap of 100 bp
*    36960    37696: contig of 737 bp in length
*    37697    37796: gap of 100 bp
*    37797    38515: contig of 719 bp in length
*    38516    38615: gap of 100 bp
*    38616    39324: contig of 709 bp in length
*    39325    39424: gap of 100 bp
*    39425    40145: contig of 721 bp in length
*    40146    40245: gap of 100 bp
*    40246    40983: contig of 738 bp in length
*    40984    41083: gap of 100 bp
*    41084    41806: contig of 723 bp in length
*    41807    41906: gap of 100 bp
*    41907    42625: contig of 719 bp in length
*    42626    42725: gap of 100 bp
*    42726    43457: contig of 732 bp in length
*    43458    43557: gap of 100 bp
*    43558    44283: contig of 726 bp in length
*    44284    44383: gap of 100 bp
*    44384    45119: contig of 736 bp in length
*    45120    45219: gap of 100 bp
*    45220    45958: contig of 739 bp in length
*    45959    46058: gap of 100 bp
*    46059    46806: contig of 748 bp in length
*    46807    46906: gap of 100 bp
*    46907    47622: contig of 716 bp in length
*    47623    47722: gap of 100 bp
*    47723    48440: contig of 718 bp in length
*    48441    48540: gap of 100 bp
*    48541    49266: contig of 726 bp in length
*    49267    49366: gap of 100 bp
*    49367    50087: contig of 721 bp in length
*    50088    50187: gap of 100 bp
*    50188    50898: contig of 711 bp in length
*    50899    50998: gap of 100 bp
*    50999    51711: contig of 713 bp in length
```

```
            *    51712    51811: gap of 100 bp
            *    51812    52482: contig of 671 bp in length
            *    52483    52582: gap of 100 bp
            *    52583    53324: contig of 742 bp in length
            *    53325    53424: gap of 100 bp
            *    53425    54150: contig of 726 bp in length
            *    54151    54250: gap of 100 bp
            *    54251    54967: contig of 717 bp in length
            *    54968    55067: gap of 100 bp
            *    55068    55780: contig of 713 bp in length
            *    55781    55880: gap of 100 bp
            *    55881    56614: contig of 734 bp in length

  Query Match             39.7%;  Score 158.6;  DB 12;  Length 68990;
  Best Local Similarity   77.7%;  Pred. No. 2.5e-43;
  Matches  206;  Conservative    1;  Mismatches   50;  Indels    8;  Gaps    1;

Qy          84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
               |||||||||||||||||| |||||||||| ||||| ||||||||||||||| ||| | |
Db       32552 CCTGTAATCCCAGCACTTTGGGAGGCCAAGCTGGGCAGATCACCTGAGGTCAGGAGTTTG 32493

Qy         144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
               | |||| |||||||||||||||||||  |||||||||||||||||||||| |||||||
Db       32492 ATACCAGCCTGGCCAACATGGTGAAACCTTGTCTTTACTAAAAATACAAAAATTAGCTGG 32433

Qy         204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
               || ||||||  ||| ||| || |||||||||||| |||| | ||| ||||||||||| ||
Db       32432 GCGTGGTGGTGCACGCCTATAATCCCAGCTACTCGGGAGGCAGAGGTTGCAGTGAGCAGA 32373

Qy         264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
               ||| |          | |   |||      | | |  | ||||||| ||:|  ||||||
Db       32372 GATTG---------TGCCATTGCACGCTAGCCTGGGCAACAAGAGTGAAACTCTATCTCAA 32321

Qy         324 AAACAACAACAAAAAACAAAAAAAC 348
               ||| || || |||||| || || ||
Db       32320 AAAAAAAAAAAAAAAAGAATAATAC 32296


RESULT 13
DD164421
LOCUS       DD164421               96593 bp    DNA     linear   PAT 23-NOV-2005
DEFINITION  NOVEL COMPOSITIONS AND METHODS FOR CANCER.
ACCESSION   DD164421
VERSION     DD164421.1  GI:83948746
KEYWORDS    JP 2005510225-A/16.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 96593)
  AUTHORS   Maurice,D.W.
  TITLE     NOVEL COMPOSITIONS AND METHODS FOR CANCER
  JOURNAL   Patent: JP 2005510225-A 16 21-APR-2005;
            SAGRES DISCOVERY
COMMENT     OS   Homo Sapien
            PN   JP 2005510225-A/16
            PD   21-APR-2005
            PF   02-DEC-2002 JP 2003546739
            PR   30-NOV-2001 US    09/997722
            PI   david w maurice
            CC
            FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .96593
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             39.6%;  Score 158.4;  DB 2;  Length 96593;
  Best Local Similarity   75.0%;  Pred. No. 3.4e-43;
  Matches  216;  Conservative    1;  Mismatches   57;  Indels   14;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||||||||||||||| |||||||| ||| ||||||||  |||||||||| |||
Db       18217 ACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATTTCCTGAGGTCAGGAGT 18276

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || ||| || |||||||||||||||||||||||||||||| ||||||||||||||||| ||||
Db       18277 TCAAGATCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 18336

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||||||||   |||||||||||||||||||| |||| | |||   | ||
Db       18337 TGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCG 18396

Qy         261 TGAGATCGCAG--------------AGTGAGCCGAAATCACAGATCACAGAGTGAGCAGA 306
                  || | |||               |||||||||| ||| ||       | ||||  ||
Db       18397 CTTGAACCCAGGCGGCGGAGGTTGCAGTGAGCCGAGATCGCACCAGCTTGGGTGACAAGC 18456

Qy         307 GTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAG 354
               |||||||: | |||||||||| ||  | |||| | |||| ||||| | |
Db       18457 GTGAGACTTCATCTCAAAAAAAAACAAAAAACAAAAAACAACCAGATG 18504


RESULT 14
AX695641
LOCUS       AX695641               96593 bp    DNA     linear   PAT 31-MAR-2003
DEFINITION  Sequence 1268 from Patent WO03008583.
ACCESSION   AX695641
VERSION     AX695641.1  GI:29418793
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Morris,D.W. and Engelhard,E.K.
  TITLE     Novel compositions and methods for cancer
  JOURNAL   Patent: WO 03008583-A 1268 30-JAN-2003;
            Sagres Discovery (US)
FEATURES             Location/Qualifiers
     source          1. .96593
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             39.6%;  Score 158.4;  DB 2;  Length 96593;
  Best Local Similarity   75.0%;  Pred. No. 3.4e-43;
  Matches  216;  Conservative    1;  Mismatches   57;  Indels   14;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||||||||||||||| |||||||| ||| ||||||||  |||||||||| |||
Db       18217 ACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATTTCCTGAGGTCAGGAGT 18276

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || ||| || |||||||||||||||||||||||||||||| ||||||||||||||||| ||||
Db       18277 TCAAGATCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 18336

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||||||||   |||||||||||||||||||| |||| | |||   | ||
Db       18337 TGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCG 18396

Qy         261 TGAGATCGCAG--------------AGTGAGCCGAAATCACAGATCACAGAGTGAGCAGA 306
                  || | |||               |||||||||| ||| ||       | ||||  ||
Db       18397 CTTGAACCCAGGCGGCGGAGGTTGCAGTGAGCCGAGATCGCACCAGCTTGGGTGACAAGC 18456

Qy         307 GTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAG 354
               |||||||: | |||||||||| ||  | |||| | |||| ||||| | |
Db       18457 GTGAGACTTCATCTCAAAAAAAAACAAAAAACAAAAAACAACCAGATG 18504


RESULT 15
AC010677/c
LOCUS       AC010677               112659 bp    DNA     linear   PRI 03-JAN-2002
DEFINITION  Homo sapiens chromosome 7 clone CTD-2304L4, complete sequence.
```

```
ACCESSION   AC010677
VERSION     AC010677.4  GI:11465112
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 112659)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 112659)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-SEP-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 112659)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-NOV-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 112659)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-2001) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 112659)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JAN-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Nov 30, 2000 this sequence version replaced gi:7630788.

            ------------------- Genome Center --------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            Contact: submissions@watson.wustl.edu
            ---------------- Project Information -----------------
            Center project name: H_MS2304L04
            ------------------------------------------------------.
FEATURES             Location/Qualifiers
     source          1. .112659
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /clone="CTD-2304L4"
ORIGIN

  Query Match             39.6%;  Score 158.4;  DB 5;  Length 112659;
  Best Local Similarity   73.9%;  Pred. No. 3.7e-43;
  Matches  201;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy         76 GCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCA 135
              | |  | |||||||||||||||||  |||||||| ||| ||| |||||||||||| ||||
Db      16959 GGTTCACGCCTGTAATCCCAGCACTCTGGGAGGCCGAGGCGGGTGGATCACCTGATGTCA 16900

Qy        136 AGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAA 195
              ||| | |||| ||  |||||||||||||||||||||||||| ||||||| |||||||||
Db      16899 GGAGTTTGAGATCAGTCTGGCCAACATGGTGAAACCCCGTCTCTACTAAATATACAAAAA 16840

Qy        196 ATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG 255
              |||||||  |||||| |||| | |||| |||||||||||| |||| |||||| ||||||
Db      16839 CCAGCTGGGTGTGGTGGTACACGCTTGTAATCCCAGCTACTCTGGAGGCGGAGGTTGCAG 16780

Qy        256 TGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKC 315
              ||||| || |  |||     |  |   | |   | | |||||| |||       |
Db      16779 TGAGCCGAAACTGCACCACTACACTCCAGCCTGGGTGACAGAGGGAGACTCTGTCTCCCA 16720

Qy        316 CGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
                 |||||  ||  ||  |||||| |||||||
Db      16719 GGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 16688


RESULT 16
AC026407
LOCUS       AC026407               155344 bp    DNA     linear   PRI 30-AUG-2001
DEFINITION  Homo sapiens chromosome 5 clone CTC-370J7, complete sequence.
ACCESSION   AC026407
VERSION     AC026407.4  GI:15375158
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 155344)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 155344)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 155344)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 155344)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Aug 30, 2001 this sequence version replaced gi:14861724.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.8% of Sequence;
            Estimated Total Number of Errors is 0.2.
            STS Content:
            WI-11894 G21380
            WI-6759 G05738
            SHGC-58349 G38490
            SHGC-57583 G37344.
FEATURES             Location/Qualifiers
     source          1. .155344
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTC-370J7"
ORIGIN

  Query Match             39.6%;  Score 158.2;  DB 5;  Length 155344;
  Best Local Similarity   78.5%;  Pred. No. 5.1e-43;
  Matches  205;  Conservative    0;  Mismatches   48;  Indels    8;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||  |||  |||||||  ||||||||||  |||
Db      95307 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAGGCGGATTACCTGAGGTCTGGAGT 95366

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || |||||| |||||||||||||||||||||| |||| ||||||||||||||||  |||
Db      95367 TCAAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATCAGC 95426

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||||||||    |||||| |||||||||||||||| ||||| ||||||||||
Db      95427 TGGGCATGGTGGCAGGGGCCTGTAATCCCAGCTACTCAGGAGGCGGAGGTTGCAGTGAGC 95486
```

```
Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||||| |          | |   |||      | | ||||  ||||  | ||| | ||||
Db      95487 TGAGATTG--------TGCCACTGCACTCCAGCCTGGGTGACAAGAGCAAAACACTGTCT 95538

Qy        321 CAAAAACAACAACAAAAAACA 341
              |||||| || || |||| |||
Db      95539 CAAAAAAAAAAAAAAAAGACA 95559
```

```
RESULT 17
AC040168
LOCUS       AC040168               170154 bp    DNA     linear   PRI 29-MAR-2003
DEFINITION  Homo sapiens chromosome 16 clone RP11-461O7, complete sequence.
ACCESSION   AC040168
VERSION     AC040168.8  GI:29366942
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 170154)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 170154)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-APR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 170154)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAY-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 170154)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JAN-2003) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   5  (bases 1 to 170154)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Mar 29, 2003 this sequence version replaced gi:27819471.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center and Los Alamos
            National Laboratory
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
FEATURES             Location/Qualifiers
     source          1. .170154
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="RP11-461O7"
ORIGIN

  Query Match             39.5%;  Score 157.8;  DB 5;  Length 170154;
  Best Local Similarity   75.8%;  Pred. No. 7.3e-43;
  Matches  207;  Conservative    1;  Mismatches   63;  Indels    2;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| |||||||||||||||| ||||||  ||||||||||||
Db     134321 ACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCAGATCAC--GAGGTCAAGAGA 134378

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||| ||||
Db     134379 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCC 134438

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||| |||||||   |||||||| ||||||||||  |||| | |||   | |
Db     134439 CGGGCGTGGTGGCTGGCACCTGTAATCCCAGCTACTTGGGAGTCTGAGGCAGGATAATCG 134498

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              || |   |   | |  |    ||||   ||     |    ||||| |||||:||||||
Db     134499 CTTGAACCTGGGAGGCGGAGGTTGCACACTCCAGCCTGGTGACAGAGAGAGACTCCGTCT 134558

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAA 353
              |||||| || || ||||||   |||||  | ||
Db     134559 CAAAAAAAAAAAAAAAAAAACCAAAAAAAAGAA 134591
```

```
RESULT 18
AC018605
LOCUS       AC018605               186747 bp    DNA     linear   HTG 24-AUG-2002
DEFINITION  Homo sapiens chromosome 16 clone RP11-764C24 map 16, WORKING DRAFT
            SEQUENCE, 15 unordered pieces.
ACCESSION   AC018605
VERSION     AC018605.4  GI:10047801
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 186747)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 16, clone RP11-764C24
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 186747)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H.,
            O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A.,
            Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
            Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A.
            and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-DEC-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 186747)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
            FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
            Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
            Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G.,
            Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
            O'Donnell,P., O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
            Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
            Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
```

```
           Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
           Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
           Zimmer,A. and Zody,M.
  TITLE    Direct Submission
  JOURNAL  Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
           Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT    On Sep 10, 2000 this sequence version replaced gi:7407949.
           All repeats were identified using RepeatMasker:
           Smit, A.F.A. & Green, P. (1996-1997)
           http://ftp.genome.washington.edu/RM/RepeatMasker.html
           -------------- Genome Center
               Center: Whitehead Institute/ MIT Center for Genome Research
               Center code: WIBR
               Web site: http://www-seq.wi.mit.edu
               Contact: sequence_submissions@genome.wi.mit.edu
           -------------- Project Information
               Center project name: L4986
               Center clone name: 764_C_24
           -------------- Summary Statistics
               Sequencing vector: M13; M77815; 100% of reads
               Chemistry: Dye-terminator Big Dye; 100% of reads
               Assembly program: Phrap; version 0.960731
               Consensus quality: 174578 bases at least Q40
               Consensus quality: 179739 bases at least Q30
               Consensus quality: 182173 bases at least Q20
               Insert size: 188000; agarose-fp
               Insert size: 185347; sum-of-contigs
               Quality coverage: 4.6 in Q20 bases; agarose-fp
               Quality coverage: 4.7 in Q20 bases; sum-of-contigs
           --------------.
           * NOTE: This is a 'working draft' sequence. It currently
           * consists of 15 contigs. The true order of the pieces
           * is not known and their order in this sequence record is
           * arbitrary. Gaps between the contigs are represented as
           * runs of N, but the exact sizes of the gaps are unknown.
           * This record will be updated with the finished sequence
           * as soon as it is available and the accession number will
           * be preserved.
           *        1    26748: contig of 26748 bp in length
           *    26749    26848: gap of 100 bp
           *    26849    28315: contig of 1467 bp in length
           *    28316    28415: gap of 100 bp
           *    28416    30899: contig of 2484 bp in length
           *    30900    30999: gap of 100 bp
           *    31000    34327: contig of 3328 bp in length
           *    34328    34427: gap of 100 bp
           *    34428    39283: contig of 4856 bp in length
           *    39284    39383: gap of 100 bp
           *    39384    44613: contig of 5230 bp in length
           *    44614    44713: gap of 100 bp
           *    44714    78567: contig of 33854 bp in length
           *    78568    78667: gap of 100 bp
           *    78668    89039: contig of 10372 bp in length
           *    89040    89139: gap of 100 bp
           *    89140    97713: contig of 8574 bp in length
           *    97714    97813: gap of 100 bp
           *    97814   110219: contig of 12406 bp in length
           *   110220   110319: gap of 100 bp
           *   110320   123656: contig of 13337 bp in length
           *   123657   123756: gap of 100 bp
           *   123757   138462: contig of 14706 bp in length
           *   138463   138562: gap of 100 bp
           *   138563   153566: contig of 15004 bp in length
           *   153567   153666: gap of 100 bp
           *   153667   180503: contig of 26837 bp in length
           *   180504   180603: gap of 100 bp
           *   180604   186747: contig of 6144 bp in length.
FEATURES             Location/Qualifiers
     source          1. .186747
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /map="16"
                     /clone="RP11-764C24"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .26748
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     gap             26749. .26848
                     /estimated_length=100
     misc_feature    26849. .28315
                     /note="assembly_fragment"
     gap             28316. .28415
                     /estimated_length=100
     misc_feature    28416. .30899
                     /note="assembly_fragment"
     gap             30900. .30999
                     /estimated_length=100
     misc_feature    31000. .34327
                     /note="assembly_fragment"
     gap             34328. .34427
                     /estimated_length=100
     misc_feature    34428. .39283
                     /note="assembly_fragment"
     gap             39284. .39383
                     /estimated_length=100
     misc_feature    39384. .44613
                     /note="assembly_fragment"
     gap             44614. .44713
                     /estimated_length=100
     misc_feature    44714. .78567
                     /note="assembly_fragment"
     gap             78568. .78667
                     /estimated_length=100
     misc_feature    78668. .89039
                     /note="assembly_fragment"
     gap             89040. .89139
                     /estimated_length=100
     misc_feature    89140. .97713
                     /note="assembly_fragment"
     gap             97714. .97813
                     /estimated_length=100
     misc_feature    97814. .110219
                     /note="assembly_fragment"
     gap             110220. .110319
                     /estimated_length=100
     misc_feature    110320. .123656
                     /note="assembly_fragment"
     gap             123657. .123756
                     /estimated_length=100
     misc_feature    123757. .138462
                     /note="assembly_fragment"
     gap             138463. .138562
                     /estimated_length=100
     misc_feature    138563. .153566
                     /note="assembly_fragment"
     gap             153567. .153666
                     /estimated_length=100
     misc_feature    153667. .180503
                     /note="assembly_fragment"
     gap             180504. .180603
                     /estimated_length=100
     misc_feature    180604. .186747
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             39.5%;  Score 157.8;  DB 12;  Length 186747;
  Best Local Similarity   75.8%;  Pred. No. 7.6e-43;
  Matches  207;  Conservative    1;  Mismatches   63;  Indels    2;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | ||||||||||||||||||| |||||||||||||||| ||||||  ||||||||||||
```

```
Db      69948 ACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCAGATCAC--GAGGTCAAGAGA 70005

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              |||||||||||||||||||||||||||||||||||| |||||||||||||||| ||||
Db      70006 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 70065

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||| |||||||   |||||||| |||||||||| |||| | |||   | |
Db      70066 CGGGCGTGGTGGCTGGCACCTGTAATCCCAGCTACTTGGGAGTCTGAGGCAGGATAATCG 70125

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              || |  |  | |  |    ||||  ||    |    ||||| |||||:||||||
Db      70126 CTTGAACCTGGGAGGCGGAGGTTGCACACTCCAGCCTGGTGACAGAGAGAGACTCCGTCT 70185

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAA 353
              |||||| || || ||||||   |||||  | ||
Db      70186 CAAAAAAAAAAAAAAAAAAACCAAAAAAAAGAA 70218
```

```
RESULT 19
AC007495
LOCUS       AC007495               198575 bp    DNA     linear   PRI 22-MAR-2003
DEFINITION  Homo sapiens chromosome 16 clone RP11-355E10, complete sequence.
ACCESSION   AC007495
VERSION     AC007495.9  GI:29150335
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 198575)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 198575)
  AUTHORS   Bruce,D., Mundt,M., Doggett,N., Munk,C., Saunders,E., Robinson,D.,
            Jones,M., Buckingham,J., Chasteen,L., Thompson,S., Goodwin,L.,
            Bryant,J., Tesmer,J., Meincke,L., Longmire,J., White,S., Tatum,O.,
            Campbell,C., Fawcett,J., Maltbie,M., Bussod,M., Sutherland,R.,
            McMurry,K., Han,C. and Deaven,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAY-1999) Center for Human Genome Studies, DOE Joint
            Genome Institute, Los Alamos National Laboratory, MS M888, Los
            Alamos, NM 87545, USA
REFERENCE   3  (bases 1 to 198575)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-FEB-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 198575)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-DEC-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   5  (bases 1 to 198575)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Mar 22, 2003 this sequence version replaced gi:27363205.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center and Los Alamos
            National Laboratory
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
            NOTE:  BACTERIAL TRANSPOSON excised at 8530.
FEATURES             Location/Qualifiers
     source          1. .198575
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="RP11-355E10"
ORIGIN
```

```
  Query Match           39.5%;  Score 157.8;  DB 5;  Length 198575;
  Best Local Similarity   75.8%;  Pred. No. 7.9e-43;
  Matches  207;  Conservative    1;  Mismatches   63;  Indels    2;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| |||||||||||||||| ||||||  ||||||||||||
Db     108130 ACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCAGATCAC--GAGGTCAAGAGA 108187

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              |||||||||||||||||||||||||||||||||||| |||||||||||||||| ||||
Db     108188 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 108247

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||| |||||||   |||||||| |||||||||| |||| | |||   | |
Db     108248 CGGGCGTGGTGGCTGGCACCTGTAATCCCAGCTACTTGGGAGTCTGAGGCAGGATAATCG 108307

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              || |  |  | |  |    ||||  ||    |    ||||| |||||:||||||
Db     108308 CTTGAACCTGGGAGGCGGAGGTTGCACACTCCAGCCTGGTGACAGAGAGAGACTCCGTCT 108367

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAA 353
              |||||| || || ||||||   |||||  | ||
Db     108368 CAAAAAAAAAAAAAAAAAAACCAAAAAAAAGAA 108400
```

```
RESULT 20
AL732374
LOCUS       AL732374               224187 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP13-444K19 on chromosome X Contains
            a mitochondrial ribosomal protein S18C (MRPS18C) pseudogene, the 3'
            end of the gene for a novel protein similar to PHD finger protein 2
            PHF2 and a CpG island, complete sequence.
ACCESSION   AL732374
VERSION     AL732374.14  GI:23476649
KEYWORDS    HTG; CpG island; MRPS18C; PHF2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 224187)
  AUTHORS   Chapman,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 2, 2002 this sequence version replaced gi:23393869.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome X, constructed by the Sanger Centre Chromosome X Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/ChrX
            RP13-444K19 is from the library RPCI-13.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
```

```
        This sequence was finished as follows unless otherwise noted: all
        regions were either double-stranded or sequenced with an alternate
        chemistry or covered by high quality data (i.e., phred quality >=
        30); an attempt was made to resolve all sequencing problems, such
        as compressions and repeats; all regions were covered by at least
        one subclone; and the assembly was confirmed by restriction digest,
        except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .224187
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP13-444K19"
                     /clone_lib="RPCI-13.2"
     gene            11617. .12039
                     /locus_tag="RP13-444K19.1-001"
                     /pseudo
     CDS             11617. .12039
                     /locus_tag="RP13-444K19.1-001"
                     /note="match: proteins: Q8R2L5 Q9Y3D5"
                     /pseudo
                     /codon_start=1
                     /product="mitochondrial ribosomal protein S18C (MRPS18C)
                     pseudogene"
     polyA_site      complement(122345)
     gene            join(complement(AL589872.9:883. .971),
                     complement(Z98051.7:6109. .6298),
                     complement(208423. .208508),complement(207924. .208032),
                     complement(203326. .203486),complement(202260. .202401),
                     complement(200042. .200228),complement(196787. .196949),
                     complement(188260. .188347),complement(187820. .187926),
                     complement(185535. .185626),complement(181358. .181447),
                     complement(179266. .179568),complement(178401. .178504),
                     complement(173431. .173609),complement(172743. .172828),
                     complement(171481. .171614),complement(170579. .170892),
                     complement(148509. .148604),complement(129799. .129908),
                     complement(125845. .126181),complement(122347. .124911))
                     /gene="PHF8"
                     /locus_tag="RP13-444K19.2-001"
     mRNA            join(complement(AL589872.9:883. .971),
                     complement(Z98051.7:6109. .6298),
                     complement(208423. .208508),complement(207924. .208032),
                     complement(203326. .203486),complement(202260. .202401),
                     complement(200042. .200228),complement(196787. .196949),
                     complement(188260. .188347),complement(187820. .187926),
                     complement(185535. .185626),complement(181358. .181447),
                     complement(179266. .179568),complement(178401. .178504),
                     complement(173431. .173609),complement(172743. .172828),
                     complement(171481. .171614),complement(170579. .170892),
                     complement(148509. .148604),complement(129799. .129908),
                     complement(125845. .126181),complement(122347. .124911))
                     /gene="PHF8"
                     /locus_tag="RP13-444K19.2-001"
                     /product="PHD finger protein 8"
                     /note="match: ESTs: AA551120.1 AL707548.1 AU118751.1
                     AW249934.1 AW304438.1 BG719138.1 BG719444.1 BG939390.1
                     BI333269.1 BI829064.1 BM462299.1 BQ025306.1 BX091300.1
                     CA314327.1
                     match: cDNAs: AB029034.1 AK022788.1"
     gene            complement(join(123382. .123724,124827. .124911,
                     125845. .126181,129799. .129908,148509. .148604,
                     170579. .170892,171481. .171614,172743. .172828,
                     173431. .173609,178401. .178504,179266. .179568,
                     181358. .181447,185535. .185626,187820. .187926,
                     188260. .188347,196787. .196949,200042. .200228,
                     202260. .202401,203326. .203486,207924. .>207927))
                     /gene="PHF8"
                     /locus_tag="RP13-444K19.2-007"
     mRNA            complement(join(123382. .123724,124827. .124911,
                     125845. .126181,129799. .129908,148509. .148604,
                     170579. .170892,171481. .171614,172743. .172828,
                     173431. .173609,178401. .178504,179266. .179568,
                     181358. .181447,185535. .185626,187820. .187926,
                     188260. .188347,196787. .196949,200042. .200228,
                     202260. .202401,203326. .203486,207924. .>207927))
                     /gene="PHF8"
                     /locus_tag="RP13-444K19.2-007"
                     /product="PHD finger protein 8"
                     /note="match: ESTs: AA620622.1 AW386032.1 AW934879.1
                     BF772185.1 BF772623.1 BG195947.1 BI828904.1
                     match: cDNAs: BC017720.1"
     CDS             complement(join(123646. .123724,124827. .124911,
                     125845. .126181,129799. .129908,148509. .148604,
                     170579. .170892,171481. .171614,172743. .172828,
                     173431. .173609,178401. .178504,179266. .179568,
                     181358. .181447,185535. .185626,187820. .187926,
                     188260. .188347,196787. .196949,200042. .200228,
                     202260. .202401,203326. .203486,207924. .>207927))
                     /gene="PHF8"
                     /locus_tag="RP13-444K19.2-007"
                     /standard_name="OTTHUMP00000023367"
                     /codon_start=3
                     /product="PHD finger protein 8"
                     /protein_id="CAI41582.1"
                     /db_xref="GI:57209883"
                     /db_xref="UniProtKB/TrEMBL:Q5JPS2"
                     /translation="SSDEVILKPTGNQLTVEFLEENSFSVPILVLKKDGLGMTLPSPS
                     FTVRDVEHYVGSDKEIDVIDVTRQADCKMKLGDFVKYYYSGKREKVLNVISLEFSDTR
                     LSNLVETPKIVRKLSWVENLWPEECVFERPNVQKYCLMSVRDSYTDFHIDFGGTSVWY
                     HVLKGEKIFYLIRPTNANLTLFECWSSSSNQNEMFFGDQVDKCYKCSVKQGQTLFIPT
                     GWIHAVLTPVDCLAFGGNFLHSLNIEMQLKAYEIEKRLSTADLFRFPNFETICWYVGK
                     HILDIFRGLRENRRHPASYLVHGGKALNLAFRAWTRKEALPDHEDEIPETVRTVQLIK
                     DLAREIRLVEDIFQQNVGKTSNIFGLQRIFPAGSIPLTRPAHSTSVSMSRLSLPSKNG
                     SKKKGLKPKELFKKAERKGKESSALGPAGQLSYNLMDTYSHQALKTGSFQKAKFNITG
                     ACLNDSDDDSPDLDLDGNESPLALLMSNGSTKRVKSLSKSRRTKIAKKVDKARLMAEQ
                     VMEDEFDLDSDDELQIDERLGKEKATLIIRPKFPRKLPRAKPCSDPNRVREPGEVEFD
                     IEEDYTTDEDMVEGVEGKLGNGSGAGGILDLLKASRQVGGPDYAALTEAPASPSTQEA
                     IQGMLCMANLQSSSSSPATSSLQAWWTGGQDRSSGSSSSGLGTVSNSPASQRTPGKRP
                     IKRPAYWRTESEEEEENASLDEQDSLGACFKDAEYIYPSLESDDDDPALKSRPKKKKN
                     SDDAPWSPKARVTPTLPKQDRPVREGTRVASIETGLAAAAAKLAQQELQKAQKKKYIK
                     KKPLLKEVEQPRPQDSNLSLTVPAPTVAATPQLVTSSSPLPPPEPKQEALSGSLADHE
                     YTARPNAFGMAQANRSTTPMAPGVFLTQRRPSVGSQSNQAGQGKRPKKGLATAKQRLG
                     RILKIHRNGKLLLRQVIVQAECRQAIHEPKLKRRDAHP"
     CDS             join(complement(AL589872.9:883. .898),
                     complement(Z98051.7:6109. .6298),
                     complement(208423. .208508),complement(207924. .208032),
                     complement(203326. .203486),complement(202260. .202401),
                     complement(200042. .200228),complement(196787. .196949),
                     complement(188260. .188347),complement(187820. .187926),
                     complement(185535. .185626),complement(181358. .181447),
                     complement(179266. .179568),complement(178401. .178504),
                     complement(173431. .173609),complement(172743. .172828),
                     complement(171481. .171614),complement(170579. .170892),
                     complement(148509. .148604),complement(129799. .129908),
                     complement(125845. .126181),complement(124823. .124911))
                     /gene="PHF8"
                     /locus_tag="RP13-444K19.2-001"
                     /standard_name="OTTHUMP00000023362"
                     /note="match: proteins: BAC65729 Q9UPP1"
                     /codon_start=1
                     /product="PHD finger protein 8"
                     /protein_id="CAI41576.1"
                     /db_xref="GI:57209878"
                     /db_xref="GOA:Q5VUJ4"
                     /db_xref="InterPro:IPR001965"
                     /db_xref="InterPro:IPR003347"
                     /db_xref="InterPro:IPR006209"
                     /db_xref="UniProtKB/TrEMBL:Q5VUJ4"
                     /translation="MNRSRAIVQRGRVLPPPAPLDTTNLAGRRTLQGRAKMASVPVYC
                     LCRLPYDVTRFMIECDMCQDWFHGSCVGVEEEKAADIDLYHCPNCEVLHGPSIMKKRR
                     GSSKGHDTHKGKPVKTGSPTFVRELRSRTFDSSDEVILKPTGNQLTVEFLEENSFSVP
                     ILVLKKDGLGMTLPSPSFTVRDVEHYVGSDKEIDVIDVTRQADCKMKLGDFVKYYYSG
                     KREKVLNVISLEFSDTRLSNLVETPKIVRKLSWVENLWPEECVFERPNVQKYCLMSVR
                     DSYTDFHIDFGGTSVWYHVLKGEKIFYLIRPTNANLTLFECWSSSSNQNEMFFGDQVD
                     KCYKCSVKQGQTLFIPTGWIHAVLTPVDCLAFGGNFLHSLNIEMQLKAYEIEKRLSTA
                     DLFRFPNFETICWYVGKHILDIFRGLRENRRHPASYLVHGGKALNLAFRAWTRKEALP
```

```
                     DHEDEIPETVRTVQLIKDLAREIRLVEDIFQQNVGKTSNIFGLQRIFPAGSIPLTRPA
                     HSTSVSMSRLSLPSKNGSKKKGLKPKELFKKAERKGKESSALGPAGQLSYNLMDTYSH
                     QALKTGSFQKAKFNITGACLNDSDDDSPDLDLDGNESPLALLMSNGSTKRVKSLSKSR
                     RTKIAKKVDKARLMAEQVMEDEFDLDSDDELQIDERLGKEKATLIIRPKFPRKLPRAK
                     PCSDPNRVREPGEVEFDIEEDYTTDEDMVEGVEGKLGNGSGAGGILDLLKASRQVGGP
                     DYAALTEAPASPSTQEAIQGMLCMANLQSSSSSPATSSLQAWWTGGQDRSSGSSSSGL
                     GTVSNSPASQRTPGKRPIKRPAYWRTESEEEEENASLDEQDSLGACFKDAEYIYPSLE
                     SDDDDPALKSRPKKKKNSDDAPWSPKARVTPTLPKQDRPVREGTRVASIETGLAAAAA
                     KLAQQELQKAQKKKYIKKKPLLKEVEQPRPQDSNLSLTVPAPTVAATPQLVTSSSPLP
                     PPEPKQEALSGSLADHEYTARPNAFGMAQANRSTTPMAPGVFLTQRRPSVGSQSNQAG
                     QGKRPKKGLATAKQRLGRILKIHRNGKLLL"
   gene              complement(join(128280. .129067,129799. .129908,
                     148509. .148604,170579. .170892,171481. .171614,
                     172743. .172828,173431. .173609,178401. .178504,
                     179266. .179568,181358. .181447,185535. .185626,

  Query Match             39.5%;  Score 157.8;  DB 5;  Length 224187;
  Best Local Similarity   70.0%;  Pred. No. 8.3e-43;
  Matches  229;  Conservative    1;  Mismatches   88;  Indels    9;  Gaps    1;

Qy         31 GAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAA 90
              ||  | ||| |   ||   |     | | | ||| |  |      |    |||||||||
Db     190819 GAGTCTAAAAAAGTATTTCAAGAAATAATGGCAGCCTGGCGCGGTGGCTCATGCCTGTAA 190878

Qy         91 TCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCAT 150
              |||||||||| |||||||||||   ||| ||||||||||||||| ||| || ||||||
Db     190879 TCCCAGCACTTTGGGAGGCCAAGGCAGGCAGATCACCTGAGGTCAGGAGTTCAAGACCAG 190938

Qy        151 CCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGT 210
              | ||||||| |||||||||||| |||| |||||||||||||||| |||| |||||||||
Db     190939 CATGGCCAATATGGTGAAACCCTGTCTGTACTAAAAATACAAAAATTAGCAGGGCATGGT 190998

Qy        211 GGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCA 270
              ||||||  |||||| |||||||| ||| |||| |||| | |||||||||| ||||||||
Db     190999 GGCACATGCCTGTAATCCCAGCTGCTCGGGAGGCGGACACTGCAGTGAGCCGAGATCGCT 191058

Qy        271 GAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAAC 330
              |  |   |          |       | |||||||| ||:|| |||||||||| ||
Db     191059 CCACTACACTCCA---------GCCTGAGCAACAGAGTGACACTCCCTCTCAAAAAAAAA 191109

Qy        331 AACAAAAAACAAAAAAACCATAAGACA 357
              ||  ||| | | ||| |  | || | |
Db     191110 AAAGAAAGAAAGAAAGAAAAGAAAAAA 191136


RESULT 21
AC150910/c
LOCUS       AC150910              169908 bp    DNA     linear   HTG 09-OCT-2004
DEFINITION  Pan troglodytes chromosome 7 clone CH251-484K21, WORKING DRAFT
            SEQUENCE, 8 unordered pieces.
ACCESSION   AC150910
VERSION     AC150910.2  GI:54019620
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 169908)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Pan troglodytes clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 169908)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-AUG-2004) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 169908)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-OCT-2004) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT     On Oct 9, 2004 this sequence version replaced gi:51241995.

            ---------- Genome Center ----------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            ---------- Project Information ----------
            Center project name: C_AB0484K21
            ----------
            ---------- Summary Statistics ----------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 164720 bases at least Q40
            Consensus quality: 165787 bases at least Q30
            Consensus quality: 166527 bases at least Q20
            ----------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 8 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1381: contig of 1381 bp in length
            *     1382     1481: gap of unknown length
            *     1482     7431: contig of 5950 bp in length
            *     7432     7531: gap of unknown length
            *     7532    21676: contig of 14145 bp in length
            *    21677    21776: gap of unknown length
            *    21777    34793: contig of 13017 bp in length
            *    34794    34893: gap of unknown length
            *    34894    54353: contig of 19460 bp in length
            *    54354    54453: gap of unknown length
            *    54454    88284: contig of 33831 bp in length
            *    88285    88384: gap of unknown length
            *    88385   125437: contig of 37053 bp in length
            *   125438   125537: gap of unknown length
            *   125538   169908: contig of 44371 bp in length.
FEATURES            Location/Qualifiers
     source         1. .169908
                    /organism="Pan troglodytes"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9598"
                    /chromosome="7"
                    /clone="CH251-484K21"
     misc_feature   1. .1381
                    /note="assembly_name:Contig31"
     gap            1382. .1481
                    /estimated_length=unknown
     misc_feature   1482. .7431
                    /note="assembly_name:Contig36
                    clone_end:SP6
                    vector_side:right"
     gap            7432. .7531
                    /estimated_length=unknown
     misc_feature   7532. .21676
                    /note="assembly_name:Contig38"
     gap            21677. .21776
                    /estimated_length=unknown
     misc_feature   21777. .34793
                    /note="assembly_name:Contig39"
     gap            34794. .34893
                    /estimated_length=unknown
     misc_feature   34894. .54353
                    /note="assembly_name:Contig40"
     gap            54354. .54453
                    /estimated_length=unknown
     misc_feature   54454. .88284
                    /note="assembly_name:Contig41"
```

```
     gap             88285. .88384
                     /estimated_length=unknown
     misc_feature    88385. .125437
                     /note="assembly_name:Contig42"
     gap             125438. .125537
                     /estimated_length=unknown
     misc_feature    125538. .169908
                     /note="assembly_name:Contig43
                     clone_end:T7
                     vector_side:left"
ORIGIN

  Query Match             39.4%;  Score 157.6;  DB 12;  Length 169908;
  Best Local Similarity   71.6%;  Pred. No. 8.6e-43;
  Matches  222;  Conservative    1;  Mismatches   80;  Indels    7;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | ||||||||| ||||||||| |||||||||||| ||| ||||||| |||||||| |||
Db      161709 ACGCCTGTAATGCCAGCACTTTGGGAGGCCAAGGCGGGAGGATCACTTGAGGTCAGGAGT 161650

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               ||||||| | |||||||||||||||||||||| |||| |||||||||||||||||| ||||
Db      161649 TCGAGACAAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGC 161590

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAG-------CCGGAGATTGC 253
               |||||||||||    |||||||||||||||||| ||  ||||       | |||||  |
Db      161589 GGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTGCTTGGGAGACTGAGCCAGGAGAATTG 161530

Qy         254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGAC 313
               |||  |   |  ||||||   ||    | |  |||||  |       |    |||||
Db      161529 CTTGAACCCGGGAGGCAGAGGTTGCAATGAGCCGAGATCGTACCACTGCAATGCAGAGAC 161470

Qy         314 KCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTT 373
               :|||||||||||| || || |||||| |||||||    ||   |  | | ||  | || |
Db      161469 TCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAGATAAAAGTAGAGACAATAGGGGGAT 161410

Qy         374 CCCAGACTAT 383
               | |||   ||
Db      161409 CTCAGCGAAT 161400


RESULT 22
AC069513
LOCUS       AC069513               174097 bp    DNA     linear   PRI 15-MAR-2003
DEFINITION  Homo sapiens 3 BAC RP11-171N2 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC069513
VERSION     AC069513.28  GI:25815276
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 174097)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
            Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N.,
            Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,
            Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
            Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
            Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
            Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E.,
            Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
            Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
            Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
            Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
            Warren,R., Washington,C., Watlington,S., Williams,G.,
            Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
            Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and
            Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 174097)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 174097)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 174097)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   5  (bases 1 to 174097)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAR-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 28, 2002 this sequence version replaced gi:23343662.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
               STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
               Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
               Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.
```

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

```
FEATURES             Location/Qualifiers
     source          1. .174097
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="RP11-171N2"
     repeat_region   60. .190
                     /rpt_family="AluSx"
     repeat_region   191. .488
                     /rpt_family="AluSc"
     repeat_region   489. .654
                     /rpt_family="AluSx"
     repeat_region   1338. .1853
                     /rpt_family="LTR40a"
     repeat_region   4196. .4483
                     /rpt_family="AluSx"
     repeat_region   4484. .4526
                     /rpt_family="FLAM"
     repeat_region   complement(5468. .5545)
                     /rpt_family="MIR"
     repeat_region   6576. .6987
                     /rpt_family="MLT1C"
     repeat_region   complement(7559. .7871)
                     /rpt_family="AluSp"
     repeat_region   8252. .8532
                     /rpt_family="AluSg"
     repeat_region   8533. .8580
                     /rpt_family="(TAAA)n"
     repeat_region   9735. .9783
                     /rpt_family="MIR"
     repeat_region   complement(13265. .13408)
                     /rpt_family="MIR"
     repeat_region   13668. .13966
                     /rpt_family="AluSx"
     repeat_region   15839. .15933
                     /rpt_family="GA-rich"
     repeat_region   15961. .16367
                     /rpt_family="(TAAA)n"
     repeat_region   16021. .16302
                     /rpt_family="AluSx"
     repeat_region   16303. .16367
                     /rpt_family="(TAAA)n"
     repeat_region   16368. .16380
                     /rpt_family="AluSx"
     repeat_region   18871. .18980
                     /rpt_family="MIR"
     repeat_region   18989. .19268
                     /rpt_family="AluSx"
     repeat_region   complement(19999. .20305)
                     /rpt_family="AluSx"
     repeat_region   complement(20696. .20770)
                     /rpt_family="MIR"
     repeat_region   complement(20877. .21305)
                     /rpt_family="L2"
     repeat_region   complement(21348. .21659)
                     /rpt_family="AluSx"
     misc_feature    21683
                     /note="AC069513:G/AC112775:A/AC124944:A"
                     /function="polymorphic site"
     repeat_region   21697. .21899
                     /rpt_family="AluSq"
     repeat_region   22331. .22442
                     /rpt_family="AluSg/x"
     repeat_region   complement(22503. .22575)
                     /rpt_family="L2"
     repeat_region   23743. .23819
                     /rpt_family="MIR"
     repeat_region   complement(24083. .24488)
                     /rpt_family="LTR38B"
     repeat_region   complement(24518. .24678)
                     /rpt_family="LTR38B"
     repeat_region   complement(25123. .25340)
                     /rpt_family="L2"
     repeat_region   26486. .26556
                     /rpt_family="MER5A"
     repeat_region   complement(26588. .26717)
                     /rpt_family="FLAM_C"
     repeat_region   complement(27202. .27497)
                     /rpt_family="AluSx"
     repeat_region   27583. .27885
                     /rpt_family="L1MC/D"
     repeat_region   27892. .28176
```

```
  Query Match          39.4%;  Score 157.4;  DB 5;  Length 174097;
  Best Local Similarity   75.8%;  Pred. No. 1e-42;
  Matches  207;  Conservative    1;  Mismatches   62;  Indels    3;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||| || |||||||| ||| ||| ||||||| |||||||| | |
Db       4214 ATGCCTGTAATCCCAGCATTTTGGGAGGCCGAGGCGGGTGGATCACTTGAGGTCAGGGGC 4273

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||| |||||||| |||||||||||||||| | ||||||||||||||||| || |
Db       4274 TTGAGACCAGCCTGGCCAGCATGGTGAAACCCCGTTTCTACTAAAAATACAAAAATTACC 4333

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||| ||||||| ||| |||||| |||||||||||||||||    |||   | ||
Db       4334 TGGGCGTGGTGGCGCACGCCTGTAATCCCAGCTACTCAGGAGGGTGAGGCAGGAGAATCG 4393

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTG---AGCAGAGTGAGACKCCG 317
                ||   | |   |    || ||||   |   |||  ||   | ||||||||||:| |
Db       4394 CTTGAATCCGGGAGGCAGAGAGATCATGCACTCCAGCCTGGGCAACAGAGTGAGACTCTG 4453

Qy        318 TCTCAAAAACAACAACAAAAAACAAAAAAACCA 350
              ||||||||| || || |||||| ||||| ||||
Db       4454 TCTCAAAAAAAAAAAAAAAAAAAAAAAAAGACCA 4486
```

```
RESULT 23
AC011495
LOCUS       AC011495               209844 bp    DNA     linear   PRI 14-JUL-2002
DEFINITION  Homo sapiens chromosome 19 clone CTB-33G10, complete sequence.
ACCESSION   AC011495
VERSION     AC011495.8  GI:21747443
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 209844)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 209844)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 209844)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (26-NOV-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 209844)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   5  (bases 1 to 209844)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 14, 2002 this sequence version replaced gi:15281207.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.6.
FEATURES             Location/Qualifiers
     source          1. .209844
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTB-33G10"
ORIGIN

  Query Match          39.4%;  Score 157.4;  DB 5;  Length 209844;
  Best Local Similarity   77.2%;  Pred. No. 1.1e-42;
  Matches  206;  Conservative    1;  Mismatches   52;  Indels    8;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| |||||||||||| ||| |||| |||||||||||| |||
Db     153894 ACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGTGGATAACCTGAGGTCAGGAGT 153953

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || |||||| ||||||||| |||||||||||||| ||| ||||||||||||||||| ||||
Db     153954 TCAAGACCAGCCTGGCCAAAATGGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 154013

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||| ||||    ||||||| |||||||||||||| |||| ||||| |||| ||||||
Db     154014 TGGGCATGATGGCGGGTACCTGTAATCCCAGCTACTCGGGAGGCGGAGGTTGCGGTGAGC 154073

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||||||           | |   |||       | | |  |  |||| || ||:| ||||
Db     154074 CGAGATCG--------TGCCATTGCACCCCAGCCTGGGCAACAAGAGCGAAACTCTGTCT 154125

Qy        321 CAAAAACAACAACAAAAAACAAAAAAA 347
              |||||| || |  ||| || ||| |||
Db     154126 CAAAAAAAAAAGAAAAGAAAAAAGAAA 154152


RESULT 24
AC074388
LOCUS       AC074388               173556 bp    DNA     linear   HTG 20-OCT-2000
DEFINITION  Homo sapiens chromosome 2 clone RP11-507P19, WORKING DRAFT
            SEQUENCE, 21 unordered pieces.
ACCESSION   AC074388
VERSION     AC074388.2  GI:9690406
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 173556)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 173556)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUL-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Aug 4, 2000 this sequence version replaced gi:9587428.

            ---------------- Genome Center ----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0507P19
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 100%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 100% of reads
            Chemistry: Dye-terminator Big Dye; 0% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 158826 bases at least Q40
            Consensus quality: 163318 bases at least Q30
            Consensus quality: 165604 bases at least Q20
            Insert size: 166000; agarose-fp
            Insert size: 171556; sum-of-contigs
            Quality coverage: 4.60 in Q20 bases; agarose-fp
            Quality coverage: 4.61 in Q20 bases; sum-of-contigs
            ----------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 21 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      2144: contig of 2144 bp in length
            *     2145      2244: gap of unknown length
            *     2245      4052: contig of 1808 bp in length
            *     4053      4152: gap of unknown length
            *     4153      6306: contig of 2154 bp in length
            *     6307      6406: gap of unknown length
            *     6407      8444: contig of 2038 bp in length
            *     8445      8544: gap of unknown length
            *     8545     11626: contig of 3082 bp in length
            *    11627     11726: gap of unknown length
            *    11727     15823: contig of 4097 bp in length
            *    15824     15923: gap of unknown length
            *    15924     20350: contig of 4427 bp in length
            *    20351     20450: gap of unknown length
            *    20451     25976: contig of 5526 bp in length
            *    25977     26076: gap of unknown length
            *    26077     31313: contig of 5237 bp in length
            *    31314     31413: gap of unknown length
            *    31414     38673: contig of 7260 bp in length
            *    38674     38773: gap of unknown length
            *    38774     45342: contig of 6569 bp in length
            *    45343     45442: gap of unknown length
            *    45443     53428: contig of 7986 bp in length
            *    53429     53528: gap of unknown length
            *    53529     60068: contig of 6540 bp in length
            *    60069     60168: gap of unknown length
            *    60169     68196: contig of 8028 bp in length
            *    68197     68296: gap of unknown length
            *    68297     76091: contig of 7795 bp in length
            *    76092     76191: gap of unknown length
            *    76192     87826: contig of 11635 bp in length
            *    87827     87926: gap of unknown length
            *    87927     99071: contig of 11145 bp in length
            *    99072     99171: gap of unknown length
            *    99172    115184: contig of 16013 bp in length
            *   115185    115284: gap of unknown length
            *   115285    132703: contig of 17419 bp in length
            *   132704    132803: gap of unknown length
            *   132804    154064: contig of 21261 bp in length
```

```
        *    154065   154164: gap of unknown length
        *    154165   173556: contig of 19392 bp in length.
FEATURES             Location/Qualifiers
    source           1. .173556
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone="RP11-507P19"
    misc_feature     1. .2144
                     /note="assembly_name:Contig11"
    gap              2145. .2244
                     /estimated_length=unknown
    misc_feature     2245. .4052
                     /note="assembly_name:Contig12"
    gap              4053. .4152
                     /estimated_length=unknown
    misc_feature     4153. .6306
                     /note="assembly_name:Contig13"
    gap              6307. .6406
                     /estimated_length=unknown
    misc_feature     6407. .8444
                     /note="assembly_name:Contig14"
    gap              8445. .8544
                     /estimated_length=unknown
    misc_feature     8545. .11626
                     /note="assembly_name:Contig15"
    gap              11627. .11726
                     /estimated_length=unknown
    misc_feature     11727. .15823
                     /note="assembly_name:Contig16"
    gap              15824. .15923
                     /estimated_length=unknown
    misc_feature     15924. .20350
                     /note="assembly_name:Contig17"
    gap              20351. .20450
                     /estimated_length=unknown
    misc_feature     20451. .25976
                     /note="assembly_name:Contig18"
    gap              25977. .26076
                     /estimated_length=unknown
    misc_feature     26077. .31313
                     /note="assembly_name:Contig19"
    gap              31314. .31413
                     /estimated_length=unknown
    misc_feature     31414. .38673
                     /note="assembly_name:Contig20"
    gap              38674. .38773
                     /estimated_length=unknown
    misc_feature     38774. .45342
                     /note="assembly_name:Contig21"
    gap              45343. .45442
                     /estimated_length=unknown
    misc_feature     45443. .53428
                     /note="assembly_name:Contig22"
    gap              53429. .53528
                     /estimated_length=unknown
    misc_feature     53529. .60068
                     /note="assembly_name:Contig23"
    gap              60069. .60168
                     /estimated_length=unknown
    misc_feature     60169. .68196
                     /note="assembly_name:Contig24"
    gap              68197. .68296
                     /estimated_length=unknown
    misc_feature     68297. .76091
                     /note="assembly_name:Contig25
                     clone_end:T7
                     vector_side:left"
    gap              76092. .76191
                     /estimated_length=unknown
    misc_feature     76192. .87826
                     /note="assembly_name:Contig26"
    gap              87827. .87926
                     /estimated_length=unknown
    misc_feature     87927. .99071
                     /note="assembly_name:Contig27"
    gap              99072. .99171
                     /estimated_length=unknown
    misc_feature     99172. .115184
                     /note="assembly_name:Contig28"
    gap              115185. .115284
                     /estimated_length=unknown
    misc_feature     115285. .132703
                     /note="assembly_name:Contig29
                     clone_end:SP6
                     vector_side:left"
    gap              132704. .132803
                     /estimated_length=unknown
    misc_feature     132804. .154064
                     /note="assembly_name:Contig30"
    gap              154065. .154164
                     /estimated_length=unknown
    misc_feature     154165. .173556
                     /note="assembly_name:Contig31"
ORIGIN

  Query Match            39.3%;  Score 157.2;  DB 12;  Length 173556;
  Best Local Similarity  74.6%;  Pred. No. 1.2e-42;
  Matches  209;  Conservative    1;  Mismatches   69;  Indels    1;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||| |||| ||||||||||||||||||||||| ||||||| |||
Db       1615 AGGCCTGTAATCCCAGTACTTTGGGAGGCCAAGGTGGGCGGATCACTTGAGGTCAGGAGT 1674

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||||||||||||| ||||||| |||| |||||||||||||||| ||||
Db       1675 TCGAGACCAGCCTGGCCAACATGGCGAAACCCTGTCTCTACTAAAAATACAAAAATTAGC 1734

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
                ||  |||||||||||||||||| | ||||||||||||||    |||   | ||
Db       1735 CAGGAGTGGTGGCACACACCTGTAATTCCAGCTACTCAGGAGGATGAGGCAGGAGAATCA 1794

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAG-CAGAGTGAGACKCCGTC 319
                 || | |||   | |  |    ||   | |  ||| || | ||  | |||||:| |||
Db       1795 CTTGAACCCAGGAGGCGGAGGTTGCAATAAGCCGAGATTGCGCCACTGCGAGACTCTGTC 1854

Qy        320 TCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATT 359
              ||||||| |  || | ||||  |||||   | ||| || |
Db    .   1855 TCAAAAAAATAAAAATAAAAATAAAAATAAAAAAGGCACT 1894


RESULT 25
AL162739/c
LOCUS       AL162739               84001 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP5-1155K23 on chromosome 1p31.3-32.3
            Contains a Down syndrome critical region gene 5 (DSCR5) pseudogene,
            a ribosomal protein S15a (RPS15A) pseudogene, a novel gene
            (FLJ10884), a ribosomal protein L36 (RPL36) pseudogene and the 3'
            end of a novel gene, complete sequence.
ACCESSION   AL162739
VERSION     AL162739.24  GI:21211651
KEYWORDS    HTG; DSCR5; FLJ10884; RPL36; RPS15A.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 84001)
  AUTHORS   White,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On May 25, 2002 this sequence version replaced gi:16973818.
            The following abbreviations are used to associate primary accession
```

```
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group.  Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RP5-1155K23 is from the library RPCI-5 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2
-------------- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vega@sanger.ac.uk
--------------
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .84001
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704K231155"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="p31.3-32.3"
                     /clone="RP5-1155K23"
                     /clone_lib="RPCI-5"
     misc_feature    2000
                     /note="Clone_right_end: RP4-537K17"
     misc_feature    4959. .4992
                     /note="Sequence from uni-directional dGTP big dye
                     terminator reads only."
     gene            join(20464. .20836,21202. .21286)
                     /locus_tag="RP5-1155K23.2-001"
                     /pseudo
     CDS             join(20464. .20836,21202. .21286)
                     /locus_tag="RP5-1155K23.2-001"
                     /note="match: proteins: Sw:P57054 Tr:AAP35702 Tr:Q9JHG1"
                     /pseudo
                     /codon_start=1
                     /product="Down syndrome critical region gene 5 (DSCR5)
                     pseudogene"
     gene            21855. .22259
                     /locus_tag="RP5-1155K23.1-001"
                     /pseudo
     CDS             21855. .22259
                     /locus_tag="RP5-1155K23.1-001"
                     /note="match: proteins: Sw:P04648 Sw:P39027 Sw:Q90YQ8
                     Tr:AAH54792 Tr:BAC56505 Tr:O77395 Tr:Q8N5M6 Tr:Q8T5T2
                     Tr:Q8UW22 Tr:Q962R3 Tr:Q9IA74"
                     /pseudo
                     /codon_start=1
                     /product="ribosomal protein S15a (RPS15A) pseudogene"
     gene            join(26164. .26254,27768. .27861,37852. .38969,41116. .43661)
                     /gene="RP5-1155K23.3"
                     /locus_tag="RP5-1155K23.3-001"
     mRNA            join(26164. .26254,27768. .27861,37852. .38969,41116. .43661)
                     /gene="RP5-1155K23.3"
                     /locus_tag="RP5-1155K23.3-001"
                     /product="novel protein (FLJ10884)"
                     /note="match: ESTs: Em:AA933744.1 Em:AA983987.1
                     Em:AI203404.1 Em:AI203902.1 Em:AI341793.1 Em:AI624743.1
                     Em:AI633279.1 Em:AI638202.1 Em:AI654816.1 Em:AI656240.1
                     Em:AI656788.1 Em:AI700108.1 Em:AU124205.1 Em:AU133378.1
                     Em:AU137168.1 Em:AU154725.1 Em:AW025668.1 Em:AW291177.1
                     Em:AW292188.1 Em:AW451461.1 Em:BF057635.1 Em:BF107193.1
                     Em:BG570847.1 Em:BQ429140.1 Em:BU677875.1 Em:BX109916.1
                     Em:BX369343.1 Em:BX382228.1 Em:BX382229.1 Em:CB988408.1
                     Em:CD642516.1 Em:CD656210.1 Em:CD656782.1
                     match: cDNAs: Em:AK001746.1 Em:AK001973.1 Em:AL834314.1
                     Em:BC012555.1"
     CDS             join(37962. .38969,41116. .42705)
                     /gene="RP5-1155K23.3"
                     /locus_tag="RP5-1155K23.3-001"
                     /standard_name="OTTHUMP00000010477"
                     /note="match: proteins: Tr:Q8NDA1 Tr:Q9NUV8 Tr:Q9NV78"
                     /codon_start=1
                     /product="novel protein (FLJ10884)"
                     /protein_id="CAI19299.1"
                     /db_xref="GI:56206204"
                     /db_xref="InterPro:IPR004244"
                     /db_xref="UniProtKB/TrEMBL:Q5T7N2"
                     /translation="MSDVSTSVQSKFARLAKKKENITYMKREQLTETDKDIAPVLDLK
                     CKDVSAIMNKFKVLMEIQDLMFEEMRETLKNDLKAVLGGKATIPEVKNSENSSSRTEF
                     QQIINLALQKTGMVGKIEGENSKIGDDNENLTFKLEVNELSGKLDNTNEYNSNDGKKL
                     PQGESRSYEVMGSMEETLCNIDDRDGNRNVHLEFTERESRKDGEDEFVKEMREERKFQ
                     KLKNKEEVLKASREEKVLMDEGAVLTLVADLSSATLDISKQWSNVFNILRENDFEPKF
                     LCEVKLAFKCDGEIKTFSDLQSLRKFASQKSSVKELLKDVLPQKEEINQGGRKYGIQE
                     KRDKTLIDSKHRAGEITSDGLSFLFLKEVKVAKPEEMKNLETQEEEFSELEELDEEAS
                     GMEDDEDTSGLEEEEEEPSGLEEEEEEEASGLEEDEASGLEEEEEQTSEQDSTFQGHT
                     LVDAKHEVEITSDGMETTFIDSVEDSESEEEEEEGKSSETGKVKTTSLTEKKASRRQKE
                     IPFSYLVGDSGKKKLVKHQVVHKTQEEEETAVPTSQGTGTPCLTLCLASPSKSLEMSH
                     DEHKKHSHTNLSISTGVTKLKKTEEKKHRTLHTEELTSKEADLTEETEENLRSSVINS
                     IREIKEEIGNLKSSHSGVLEIENSVDDLSSRMDILEERIDSLEDQIEEFSKDTMQMTK
                     QIISKERQRDIEERSRSCNIRLIGIPEKESYENRAEDIIKEIIDENFAELKKGSSLEI
                     VSACRVPSKIDEKRLTPRHILVKFWNSSDKEKIIRASRERREITYQGTRIRLTADLSL
                     DTLDARSKWSNVFKVLLEKGFNPRILYPAKMAFDFRGKTKVFLSIEEFRDYVLHMPTL
                     RELLGNNIP"
     gene            complement(39469. .39775)
                     /locus_tag="RP5-1155K23.4-001"
                     /pseudo
     CDS             complement(39469. .39775)
                     /locus_tag="RP5-1155K23.4-001"
                     /note="match: proteins: Sw:P39032 Sw:P47964 Sw:Q9Y3U8
                     Tr:AAH21595 Tr:BAB22575 Tr:Q8UW19 Tr:Q90YT2 Tr:Q98TF6"
                     /pseudo
                     /codon_start=1
                     /product="ribosomal protein L36 (RPL36) pseudogene"
     polyA_signal    42864. .42869
                     /gene="RP5-1155K23.3"
                     /locus_tag="RP5-1155K23.3-001"
     polyA_site      42887
                     /gene="RP5-1155K23.3"
                     /locus_tag="RP5-1155K23.3-001"
     polyA_signal    43636. .43641
                     /gene="RP5-1155K23.3"
                     /locus_tag="RP5-1155K23.3-001"
     polyA_site      43661
                     /gene="RP5-1155K23.3"
                     /locus_tag="RP5-1155K23.3-001"
     gene            join(complement(AL139343.9:68437. .68745),
                     complement(AL139343.9:30881. .30966),
                     complement(AL139343.9:22536. .24419),
                     complement(AL139343.9:20810. .20921),
                     complement(AL139343.9:17619. .17837),
                     complement(AL139343.9:16064. .16151),
                     complement(AL139343.9:12424. .12643),
                     complement(AL139343.9:2399. .2541),
                     complement(78805. .79005),complement(68312. .69714))
                     /gene="RP5-1155K23.5"
                     /locus_tag="RP5-1155K23.5-001"
     gene            join(complement(AL139343.9:12424. .12756),
                     complement(AL139343.9:2399. .2541),
                     complement(78805. .79005),complement(68312. .69714))
                     /gene="RP5-1155K23.5"
                     /locus_tag="RP5-1155K23.5-004"
     mRNA            join(complement(AL139343.9:12424. .12756),
                     complement(AL139343.9:2399. .2541),
```

```
                     complement(78805. .79005),complement(68312. .69714))
                     /gene="RP5-1155K23.5"
                     /locus_tag="RP5-1155K23.5-001"
                     /product="novel protein"
                     /note="match: ESTs: Em:AW409800.1 Em:BE019514.1
                     Em:BE295315.1 Em:BE300309.1 Em:BF204822.1 Em:BF304949.1
                     Em:BG825977.1
                     match: cDNAs: Em:AK096259.1 Em:BC033147.1"
     mRNA            join(complement(AL139343.9:68437. .68745),
                     complement(AL139343.9:30881. .30966),
                     complement(AL139343.9:22536. .24419),
                     complement(AL139343.9:20810. .20921),
                     complement(AL139343.9:17619. .17837),
                     complement(AL139343.9:16064. .16151),
                     complement(AL139343.9:12424. .12643),
                     complement(AL139343.9:2399. .2541),
                     complement(78805. .79005),complement(68312. .69714))
                     /gene="RP5-1155K23.5"
                     /locus_tag="RP5-1155K23.5-001"
                     /product="novel protein"
                     /note="match: ESTs: Em:AA358454.1 Em:AA456955.1
                     Em:AA457064.1 Em:AA680112.1 Em:AI333316.1 Em:AI381634.1
                     Em:AI420038.1 Em:AI468658.1 Em:AI741903.1 Em:AI970049.1
                     Em:AV684541.1 Em:AV684569.1 Em:AV687501.1 Em:AV688590.1
                     Em:AV688594.1 Em:AV689780.1 Em:AV694201.1 Em:AW020342.1
                     Em:AW023699.1 Em:AW955294.1 Em:BE162719.1 Em:BE206280.1
                     Em:BE299669.1 Em:BE300482.1 Em:BE537560.1 Em:BG825977.1
                     Em:BI090563.1 Em:BI491401.1 Em:BM698610.1 Em:BM701638.1
                     Em:BQ022752.1 Em:BQ025863.1 Em:BU738703.1 Em:BX113154.1
                     Em:CB960340.1 Em:CB960708.1 Em:CB960937.1 Em:CB960938.1
                     Em:CB989254.1 Em:CB990362.1 Em:CD514723.1 Em:N55540.1
                     Em:R20871.1 Em:R42051.1

 Query Match             39.3%;  Score 157;  DB 5;  Length 84001;
 Best Local Similarity   77.4%;  Pred. No. 1e-42;
 Matches  205;  Conservative    1;  Mismatches   51;  Indels    8;  Gaps    1;

Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
              |||||||||||||||||| |||||||||| ||||| |||||||||||||| ||| | |
Db      57715 CCTGTAATCCCAGCACTTTGGGAGGCCAAGCTGGGCAGATCACCTGAGGTCAGGAGTTTG 57656

Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
              | |||| ||||||||||||||||||||||  |||| |||||||||||||||| |||||||
Db      57655 ATACCAGCCTGGCCAACATGGTGAAACCTTGTCTCTACTAAAAATACAAAAATTAGCTGG 57596

Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
              || ||||||  ||| ||| || ||||||||||||| |||| | ||| ||||||||||| ||
Db      57595 GCGTGGTGGTGCACGCCTATAATCCCAGCTACTCGGGAGGCAGAGGTTGCAGTGAGCAGA 57536

Qy        264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
              ||| |          | |   |||      | | |  | ||||||| ||:|  ||||||
Db      57535 GATTG--------TGCCATTGCACGCTAGCCTGGGCAACAAGAGTGAAACTCTATCTCAA 57484

Qy        324 AAACAACAACAAAAAACAAAAAAAC 348
              ||| || || |||||| || || ||
Db      57483 AAAAAAAAAAAAAAAAGAATAATAC 57459


RESULT 26
AC025599/c
LOCUS       AC025599               158330 bp    DNA     linear   PRI 01-FEB-2002
DEFINITION  Homo sapiens chromosome 8, clone RP11-508K19, complete sequence.
ACCESSION   AC025599
VERSION     AC025599.8  GI:18464214
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 158330)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 8, clone RP11-508K19
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 158330)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 158330)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-FEB-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 1, 2002 this sequence version replaced gi:18072197.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L5714
                Center clone name: 508_K_19
            ----------.
FEATURES              Location/Qualifiers
     source           1. .158330
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="8"
                      /map="8"
                      /clone="RP11-508K19"
```

```
                /clone_lib="RPCI-11 Human Male BAC"
   unsure       336. .346
                /note="<30 qual SNGL region."
   repeat_region 378. .809
                /rpt_family="MLT2A1"
   repeat_region complement(1447. .1505)
                /rpt_family="MIR"
   repeat_region complement(1838. .2106)
                /rpt_family="L2"
   repeat_region complement(2113. .2274)
                /rpt_family="L2"
   repeat_region 3269. .3311
                /rpt_family="AT_rich"
   repeat_region 3802. .4110
                /rpt_family="AluY"
   repeat_region 4499. .4560
                /rpt_family="GA-rich"
   repeat_region 4679. .4709
                /rpt_family="(TA)n"
   repeat_region 4995. .5034
                /rpt_family="(TAAAAA)n"
   repeat_region complement(5268. .5423)
                /rpt_family="MER3"
   repeat_region complement(5912. .6354)
                /rpt_family="L1MA4"
   repeat_region 6458. .6542
                /rpt_family="MIR"
   repeat_region 7696. .7847
                /rpt_family="FRAM"
   repeat_region 8274. .8987
                /rpt_family="L1M4"
   repeat_region 8990. .9123
                /rpt_family="MIR"
   repeat_region 9545. .9907
                /rpt_family="THE1B"
   repeat_region 10607. .10959
                /rpt_family="L1M4"
   repeat_region 11222. .11285
                /rpt_family="(TTATA)n"
   repeat_region 11286. .11403
                /rpt_family="AT_rich"
   repeat_region complement(11709. .12080)
                /rpt_family="L1PB3"
   repeat_region complement(12081. .12260)
                /rpt_family="AluSx"
   repeat_region 12261. .12287
                /rpt_family="(TTTG)n"
   repeat_region complement(12288. .12409)
                /rpt_family="AluSx"
   repeat_region complement(12437. .13784)
                /rpt_family="L1PA6"
   unsure       13197. .13202
                /note="single clone coverage."
   unsure       13227. .13236
                /note="<30 qual SNGL region"
   unsure       13266. .13273
                /note="<30 qual SNGL region"
   unsure       13311. .13322
                /note="<30 qual SNGL region"
   unsure       13336. .13340
                /note="<30 qual SNGL region"
   repeat_region complement(13785. .13893)
                /rpt_family="L1PB3"
   repeat_region complement(13895. .14264)
                /rpt_family="Tigger2a"
   repeat_region 14270. .14377
                /rpt_family="L1MC/D"
   repeat_region 15674. .15729
                /rpt_family="GA-rich"
   repeat_region complement(15921. .16227)
                /rpt_family="AluSx"
   repeat_region complement(16455. .16545)
                /rpt_family="MER69"
   repeat_region 17045. .17462
                /rpt_family="MSTB"
   repeat_region complement(18079. .18366)
                /rpt_family="LTR45"
   repeat_region complement(18479. .18774)
                /rpt_family="AluSx"
   repeat_region 19235. .19484
                /rpt_family="L1MC4a"
   repeat_region 21909. .22067
                /rpt_family="MIR"
   repeat_region 22107. .22141
                /rpt_family="(CA)n"
   repeat_region complement(23108. .23256)
                /rpt_family="MIR"
   repeat_region 23382. .23758
                /rpt_family="THE1C"
   repeat_region complement(24728. .24964)
                /rpt_family="L1MC/D"
   unsure       24760. .24783
                /note="<30 qual SNGl region"
   repeat_region complement(25032. .25685)
                /rpt_family="L1MC/D"
   repeat_region 25668. .26379
                /rpt_family="L1MC/D"
   repeat_region complement(26384. .26656)
                /rpt_family="MLT1C"
   repeat_region 26691. .26898
                /rpt_family="L1ME1"
   repeat_region 26899. .26986
                /rpt_family="LTR37B"
   repeat_region 27060. .27467
                /rpt_family="L1ME1"
   repeat_region complement(27468. .27777)
                /rpt_family="AluY"
   repeat_region 27778. .27839
                /rpt_family="L1ME1"
   repeat_region complement(27881. .34258)
                /rpt_family="L1PA7"
   repeat_region 34351. .34381
                /rpt_family="AT_rich"
   repeat_region complement(35666. .36486)

  Query Match             39.2%;  Score 156.8;  DB 5;  Length 158330;
  Best Local Similarity   71.0%;  Pred. No. 1.6e-42;
  Matches  206;  Conservative    1;  Mismatches   83;  Indels    0;  Gaps    0;

Qy          58 AGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGG 117
               || || ||| ·| | | |     ||||||||||||||||||||| |||||||  ||| ||
Db       71945 AGCCCTGGCCAGGCGCGGTACTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCGG 71886

Qy         118 GCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCT 177
               | |||||||||||| ||| ||| || |||||| |||||||||||||  |||||||| ||
Db       71885 GTGGATCACCTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGCAAAACCCCGGCT 71826

Qy         178 TTACTAAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTC 237
               ||||||||| ||||||| ||||  ||| | ||||||  |||||||| |||||| ||| |
Db       71825 CTACTAAAAAGACAAAAACTAGCCAGGCTTTGTGGCAGGCACCTGTAATCCCAGTTACCC 71766

Qy         238 AGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGA 297
               ||||| | |||   |     |    ||   || |||   ||   | |         |
Db       71765 AGGAGGCTGAGGCAGGGAGAATTACAGTGAGCCGAGATCGCACCACTAGACTCCAGCCTG 71706

Qy         298 GTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
               |  | |||||||||||:|||||||||||  || || |||||| ||| |||
Db       71705 GGCAACAGAGTGAGACTCCGTCTCAAAATAAAAAATAAAAAATAAATAAA 71656


RESULT 27
AC068995/c
LOCUS       AC068995              195070 bp    DNA     linear   HTG 31-DEC-2000
DEFINITION  Homo sapiens chromosome 3 clone RP11-463G10, WORKING DRAFT
            SEQUENCE, 23 unordered pieces.
ACCESSION   AC068995
```

```
VERSION     AC068995.10  GI:11995516
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 195070)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,D., Newtson,J., Newtson,N.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
            Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
            Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
            Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
            Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N.,
            Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
            Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
            Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
            Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.
            and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 195070)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Dec 29, 2000 this sequence version replaced gi:11024742.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: HBGH
                Center clone name: RP11-463G10
            -------------- Summary Statistics
                Sequencing vector: M13; L08821
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 160818 bases at least Q40
                Consensus quality: 177694 bases at least Q30
                Consensus quality: 184433 bases at least Q20
                Estimated insert size: 185987; sum-of-contigs estimation
                Quality coverage: 0x in Q20 bases; agarose-fp estimation
                Quality coverage: 3.4x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 23 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    20158: contig of 20158 bp in length
            *    20159    20258: gap of unknown length
            *    20259    32636: contig of 12378 bp in length
            *    32637    32736: gap of unknown length
            *    32737    49656: contig of 16920 bp in length
            *    49657    49756: gap of unknown length
            *    49757    63725: contig of 13969 bp in length
            *    63726    63825: gap of unknown length
            *    63826    78734: contig of 14909 bp in length
            *    78735    78834: gap of unknown length
            *    78835    95253: contig of 16419 bp in length
            *    95254    95353: gap of unknown length
            *    95354   108563: contig of 13210 bp in length
            *   108564   108663: gap of unknown length
            *   108664   120335: contig of 11672 bp in length
            *   120336   120435: gap of unknown length
            *   120436   130876: contig of 10441 bp in length
            *   130877   130976: gap of unknown length
            *   130977   139482: contig of 8506 bp in length
            *   139483   139582: gap of unknown length
            *   139583   147337: contig of 7755 bp in length
            *   147338   147437: gap of unknown length
            *   147438   155046: contig of 7609 bp in length
            *   155047   155146: gap of unknown length
            *   155147   162104: contig of 6958 bp in length
            *   162105   162204: gap of unknown length
            *   162205   168086: contig of 5882 bp in length
            *   168087   168186: gap of unknown length
            *   168187   175732: contig of 7546 bp in length
            *   175733   175832: gap of unknown length
            *   175833   180404: contig of 4572 bp in length
            *   180405   180504: gap of unknown length
            *   180505   182399: contig of 1895 bp in length
            *   182400   182499: gap of unknown length
            *   182500   184812: contig of 2313 bp in length
            *   184813   184912: gap of unknown length
            *   184913   187927: contig of 3015 bp in length
            *   187928   188027: gap of unknown length
            *   188028   189826: contig of 1799 bp in length
            *   189827   189926: gap of unknown length
            *   189927   192637: contig of 2711 bp in length
            *   192638   192737: gap of unknown length
            *   192738   193873: contig of 1136 bp in length
            *   193874   193973: gap of unknown length
            *   193974   195070: contig of 1097 bp in length.
FEATURES             Location/Qualifiers
     source          1. .195070
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="RP11-463G10"
     gap             20159. .20258
                     /estimated_length=unknown
     gap             32637. .32736
                     /estimated_length=unknown
     gap             49657. .49756
                     /estimated_length=unknown
     gap             63726. .63825
                     /estimated_length=unknown
     gap             78735. .78834
```

```
                     /estimated_length=unknown
     gap             95254. .95353
                     /estimated_length=unknown
     gap             108564. .108663
                     /estimated_length=unknown
     gap             120336. .120435
                     /estimated_length=unknown
     gap             130877. .130976
                     /estimated_length=unknown
     gap             139483. .139582
                     /estimated_length=unknown
     gap             147338. .147437
                     /estimated_length=unknown
     gap             155047. .155146
                     /estimated_length=unknown
     gap             162105. .162204
                     /estimated_length=unknown
     gap             168087. .168186
                     /estimated_length=unknown
     gap             175733. .175832
                     /estimated_length=unknown
     gap             180405. .180504
                     /estimated_length=unknown
     gap             182400. .182499
                     /estimated_length=unknown
     gap             184813. .184912
                     /estimated_length=unknown
     gap             187928. .188027
                     /estimated_length=unknown
     gap             189827. .189926
                     /estimated_length=unknown
     gap             192638. .192737
                     /estimated_length=unknown
     gap             193874. .193973
                     /estimated_length=unknown
ORIGIN

  Query Match             39.2%;  Score 156.8;  DB 12;  Length 195070;
  Best Local Similarity   71.0%;  Pred. No. 1.8e-42;
  Matches  206;  Conservative    1;  Mismatches   83;  Indels    0;  Gaps    0;

Qy          58 AGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGG 117
               || || |||  | | | |     ||||||||||||||||||||||| |||||||  ||| ||
Db        2493 AGCCCTGGCCAGGCGCGGTACTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCGG 2434

Qy         118 GCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCT 177
               | |||||||||||| ||| ||| || |||||| ||||||||||||||  |||||||| ||
Db        2433 GTGGATCACCTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGCAAAACCCCGGCT 2374

Qy         178 TTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTC 237
               ||||||||| ||||||| ||||  ||| | ||||||  |||||||| |||||| ||| |
Db        2373 CTACTAAAAAGACAAAAACTAGCCAGGCTTTGTGGCAGGCACCTGTAATCCCAGTTACCC 2314

Qy         238 AGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGA 297
               ||||| | |||   |     |    ||   || |||   ||   | |          |
Db        2313 AGGAGGCTGAGGCAGGGAGAATTACAGTGAGCCGAGATCGCACCACTAGACTCCAGCCTG 2254

Qy         298 GTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
               | | ||||||||||:|||||||||||  || || |||||| ||| |||
Db        2253 GGCAACAGAGTGAGACTCCGTCTCAAAATAAAAAATAAAAAATAAATAAA 2204


RESULT 28
AL353145/c
LOCUS       AL353145              111372 bp    DNA     linear   HTG 10-JUL-2001
DEFINITION  Homo sapiens chromosome 1 clone RP4-633K13 map p34.3-36.11, 9
            unordered pieces.
ACCESSION   AL353145
VERSION     AL353145.4  GI:9796994
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mclay,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 12, 2000 this sequence version replaced gi:9213061.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: dJ633K13
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator ABI; 1% of reads
            Chemistry: Dye-terminator Big Dye; 74% of reads
            Chemistry: Dye-terminator ET-amersham; 24% of reads Consensus
            quality: 105338 bases at least Q40
            Consensus quality: 107404 bases at least Q30
            Consensus quality: 108616 bases at least Q20
            Insert size: 110572; sum-of-contigs
            Insert size: 124958; 7.7% error; agarose-fp
            Quality coverage: 3.70x in Q20 bases; sum-of-contigs Quality
            coverage: 3.47x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 9 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     3034: contig of 3034 bp in length
            *     3035     3134: gap of 100 bp
            *     3135    12408: contig of 9274 bp in length
            *    12409    12508: gap of 100 bp
            *    12509    31917: contig of 19409 bp in length
            *    31918    32017: gap of 100 bp
            *    32018    49801: contig of 17784 bp in length
            *    49802    49901: gap of 100 bp
            *    49902    75093: contig of 25192 bp in length
            *    75094    75193: gap of 100 bp
            *    75194    80482: contig of 5289 bp in length
            *    80483    80582: gap of 100 bp
            *    80583    93069: contig of 12487 bp in length
            *    93070    93169: gap of 100 bp
            *    93170   105801: contig of 12632 bp in length
            *   105802   105901: gap of 100 bp
            *   105902   111372: contig of 5471 bp in length.
FEATURES             Location/Qualifiers
     source          1. .111372
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="p34.3-36.11"
                     /clone="RP4-633K13"
                     /clone_lib="RPCI-4"
     misc_feature    1. .3034
                     /note="assembly_fragment:00063
                     fragment_chain:1"
     misc_feature    3135. .12408
                     /note="assembly_fragment:01010
                     fragment_chain:1"
     misc_feature    12509. .31917
                     /note="assembly_fragment:01234
                     fragment_chain:1"
```

```
     misc_feature    32018. .49801
                     /note="assembly_fragment:00425
                     fragment_chain:2"
     misc_feature    49902. .75093
                     /note="assembly_fragment:01051
                     fragment_chain:2"
     misc_feature    75194. .80482
                     /note="assembly_fragment:01372
                     fragment_chain:3"
     misc_feature    80583. .93069
                     /note="assembly_fragment:00629
                     fragment_chain:3"
     misc_feature    93170. .105801
                     /note="assembly_fragment:00500"
     misc_feature    105902. .111372
                     /note="assembly_fragment:00583"
ORIGIN

  Query Match             39.2%;  Score 156.6;  DB 12;  Length 111372;
  Best Local Similarity   71.1%;  Pred. No. 1.6e-42;
  Matches  207;  Conservative    0;  Mismatches   84;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               ||||||||||||||||||||| |||||||  |||  || ||  |||||||||||||  ||
Db      104784 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAGGTGGGTCACCTGAGGTCAGAAGT 104725

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               | ||||||| |||||||||||||||||||||||||| ||||||||||||||||  ||
Db      104724 TTGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATCAGT 104665

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||||||||||  | |||||| || |||||||| |||| | ||| | | ||
Db      104664 CGGGCATGGTGGCAGGCGCCTGTAATCTCAGCTACTCGGGAGGCTGAGGTGGGAGAATTG 104605

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                 || | |||||   || |  | |  ||||| ||   |  |  ||  |||      ||
Db      104604 CTTGACCCCAGAGGTTGCAGTGAGCTGAGATCGCACCATTGGGCGACAGAGCAAGTCTCC 104545

Qy         321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGG 371
               ||  | ||||| | |||||||||| ||  |||  |     |  | ||| ||
Db      104544 AAATCAAAACAAAACAAAACAAAACAAATATATTAGGAATTAAAGCTGTGG 104494


RESULT 29
AC022252
LOCUS       AC022252               150934 bp    DNA     linear   HTG 16-MAR-2000
DEFINITION  Homo sapiens clone RP11-28O12, WORKING DRAFT SEQUENCE, 28 unordered
            pieces.
ACCESSION   AC022252
VERSION     AC022252.2  GI:7249106
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 150934)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-28O12
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 150934)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 16, 2000 this sequence version replaced gi:6778512.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L4795
                Center clone name: 28_O_12
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 132159 bases at least Q40
                Consensus quality: 141146 bases at least Q30
                Consensus quality: 144793 bases at least Q20
                Insert size: 157000; agarose-fp
                Insert size: 148234; sum-of-contigs
                Quality coverage: 3.1 in Q20 bases; agarose-fp
                Quality coverage: 3.2 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 28 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1563: contig of 1563 bp in length
            *     1564     1663: gap of 100 bp
            *     1664     2811: contig of 1148 bp in length
            *     2812     2911: gap of 100 bp
            *     2912     4293: contig of 1382 bp in length
            *     4294     4393: gap of 100 bp
            *     4394     4757: contig of 364 bp in length
            *     4758     4857: gap of 100 bp
            *     4858     6513: contig of 1656 bp in length
            *     6514     6613: gap of 100 bp
            *     6614     8468: contig of 1855 bp in length
            *     8469     8568: gap of 100 bp
            *     8569    10457: contig of 1889 bp in length
            *    10458    10557: gap of 100 bp
            *    10558    12036: contig of 1479 bp in length
            *    12037    12136: gap of 100 bp
            *    12137    13936: contig of 1800 bp in length
            *    13937    14036: gap of 100 bp
            *    14037    16726: contig of 2690 bp in length
            *    16727    16826: gap of 100 bp
            *    16827    20428: contig of 3602 bp in length
            *    20429    20528: gap of 100 bp
            *    20529    24488: contig of 3960 bp in length
            *    24489    24588: gap of 100 bp
            *    24589    28867: contig of 4279 bp in length
            *    28868    28967: gap of 100 bp
            *    28968    33307: contig of 4340 bp in length
            *    33308    33407: gap of 100 bp
            *    33408    35317: contig of 1910 bp in length
            *    35318    35417: gap of 100 bp
            *    35418    39417: contig of 4000 bp in length
            *    39418    39517: gap of 100 bp
            *    39518    45499: contig of 5982 bp in length
```

```
            *    45500    45599: gap of 100 bp
            *    45600    50734: contig of 5135 bp in length
            *    50735    50834: gap of 100 bp
            *    50835    57226: contig of 6392 bp in length
            *    57227    57326: gap of 100 bp
            *    57327    64154: contig of 6828 bp in length
            *    64155    64254: gap of 100 bp
            *    64255    69806: contig of 5552 bp in length
            *    69807    69906: gap of 100 bp
            *    69907    77996: contig of 8090 bp in length
            *    77997    78096: gap of 100 bp
            *    78097    88132: contig of 10036 bp in length
            *    88133    88232: gap of 100 bp
            *    88233    98038: contig of 9806 bp in length
            *    98039    98138: gap of 100 bp
            *    98139   109294: contig of 11156 bp in length
            *   109295   109394: gap of 100 bp
            *   109395   123250: contig of 13856 bp in length
            *   123251   123350: gap of 100 bp
            *   123351   134659: contig of 11309 bp in length
            *   134660   134759: gap of 100 bp
            *   134760   150934: contig of 16175 bp in length.
FEATURES             Location/Qualifiers
     source          1. .150934
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-28O12"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1563
                     /note="assembly_fragment"
     gap             1564. .1663
                     /estimated_length=100
     misc_feature    1664. .2811
                     /note="assembly_fragment"
     gap             2812. .2911
                     /estimated_length=100
     misc_feature    2912. .4293
                     /note="assembly_fragment"
     gap             4294. .4393
                     /estimated_length=100
     misc_feature    4394. .4757
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     gap             4758. .4857
                     /estimated_length=100
     misc_feature    4858. .6513
                     /note="assembly_fragment"
     gap             6514. .6613
                     /estimated_length=100
     misc_feature    6614. .8468
                     /note="assembly_fragment"
     gap             8469. .8568
                     /estimated_length=100
     misc_feature    8569. .10457
                     /note="assembly_fragment"
     gap             10458. .10557
                     /estimated_length=100
     misc_feature    10558. .12036
                     /note="assembly_fragment"
     gap             12037. .12136
                     /estimated_length=100
     misc_feature    12137. .13936
                     /note="assembly_fragment"
     gap             13937. .14036
                     /estimated_length=100
     misc_feature    14037. .16726
                     /note="assembly_fragment"
     gap             16727. .16826
                     /estimated_length=100
     misc_feature    16827. .20428
                     /note="assembly_fragment"
     gap             20429. .20528
                     /estimated_length=100
     misc_feature    20529. .24488
                     /note="assembly_fragment"
     gap             24489. .24588
                     /estimated_length=100
     misc_feature    24589. .28867
                     /note="assembly_fragment"
     gap             28868. .28967
                     /estimated_length=100
     misc_feature    28968. .33307
                     /note="assembly_fragment"
     gap             33308. .33407
                     /estimated_length=100
     misc_feature    33408. .35317
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     gap             35318. .35417
                     /estimated_length=100
     misc_feature    35418. .39417
                     /note="assembly_fragment"
     gap             39418. .39517
                     /estimated_length=100
     misc_feature    39518. .45499
                     /note="assembly_fragment"

  Query Match             39.2%;  Score 156.6;  DB 12;  Length 150934;
  Best Local Similarity   73.1%;  Pred. No. 1.8e-42;
  Matches  201;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||||| ||||||||||||  |||||||||||||||||||| |||
Db      87071 ATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCAGGCGGATCACCTGAGGTCAGGAGT 87130

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              || |||||| |||| |||||||||||||||||||| ||||||||||||||| ||||
Db      87131 TCAAGACCAGCCTGACCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 87190

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||  |||||||||||| |||||||||||||||||||||| | |||  || ||
Db      87191 TGGGTGTGGTGGCACACGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGATGGAGAATCT 87250

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                || | | |    |    ||   ||  || |  ||| || ||      |        |
Db      87251 CTTGAACCCGGGAGGTAGAGACTGCAGTGAGCCGAGACTGCGCCACTGCATGACAGAGCG 87310

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGA 355
              | | || |  |||||| ||||||  | | ||
Db      87311 AGACACCATCTCAAAAAAAAAAAAAAAAAAAAGGA 87345


RESULT 30
AC108040/c
LOCUS       AC108040               163521 bp    DNA     linear   HTG 30-JAN-2002
DEFINITION  Homo sapiens chromosome 4 clone RP11-210O10, WORKING DRAFT
            SEQUENCE, 4 unordered pieces.
ACCESSION   AC108040
VERSION     AC108040.2  GI:18425316
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 163521)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 163521)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JAN-2002) Genome Sequencing Center, Washington
```

```
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jan 30, 2002 this sequence version replaced gi:18308830.

            ------------------- Genome Center --------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            Contact: submissions@watson.wustl.edu
            ---------------- Project Information -----------------
            Center project name: H_NH0210O10
            ------------------------------------------------------
            ------------- Summary Statistics  -------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 162304 bases at least Q40
            Consensus quality: 162444 bases at least Q30
            Consensus quality: 162529 bases at least Q20
            Insert size: 189000; agarose-fp
            Insert size: 163221; sum-of-contigs
            Quality coverage: 8.88 in Q20 bases; agarose-fp
            Quality coverage: 8.85 in Q20 bases; sum-of-contigs
            ------------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1564: contig of 1564 bp in length
            *     1565     1664: gap of unknown length
            *     1665    88984: contig of 87320 bp in length
            *    88985    89084: gap of unknown length
            *    89085   112702: contig of 23618 bp in length
            *   112703   112802: gap of unknown length
            *   112803   163521: contig of 50719 bp in length.
FEATURES                Location/Qualifiers
    source              1. .163521
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="4"
                        /clone="RP11-210O10"
    misc_feature        1. .1564
                        /note="assembly_name:Contig25
                        clone_end:SP6
                        vector_side:left"
    gap                 1565. .1664
                        /estimated_length=unknown
    misc_feature        1665. .88984
                        /note="assembly_name:Contig28
                        clone_end:T7
                        vector_side:left"
    gap                 88985. .89084
                        /estimated_length=unknown
    misc_feature        89085. .112702
                        /note="assembly_name:Contig26"
    gap                 112703. .112802
                        /estimated_length=unknown
    misc_feature        112803. .163521
                        /note="assembly_name:Contig27"
ORIGIN

  Query Match             39.2%;  Score 156.6;  DB 12;  Length 163521;
  Best Local Similarity   71.4%;  Pred. No. 1.9e-42;
  Matches  222;  Conservative    1;  Mismatches   80;  Indels    8;  Gaps    1;

Qy         37 AAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAG 96
                |||  ||  |||  | ||    | ||  |    |       |    |||| |||||||| ||
Db      98071 AAAAATATATACAAAATACAGAAGCTGGCTGGGTGTGATGGCTCATGCTTGTAATCCTAG 98012

Qy         97 CACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGC 156
               ||||| |||||||||||||  || ||||||||||||||| ||| ||||||| | ||||||
Db      98011 CACTTTGGGAGGCCAAGGCAGGTGGATCACCTGAGGTCAGGAGTTCGAGACTAGCCTGGC 97952

Qy        157 CAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACA 216
              |||||||||||||||||||| |||||||||||||| | ||||||||| |||||||
Db      97951 CAACATGGTGAAACCCCGTCTCTACTAAAAATACAAACATTAGCTGGGCTTGGTGGCGGG 97892

Qy        217 CACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGA 276
              | | |||| |||||||||||| ||||| | ||| ||||||||||||  |||||        |
Db      97891 CGCTTGTAATCCCAGCTACTGAGGAGGCAGAGGTTGCAGTGAGCCTAGATC--------A 97840

Qy        277 GCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAA 336
              | |   |||       | | ||  ||||||| ||:|| |||||||||  || ||  ||
Db      97839 TGCCATTGCACTCTGGCCTGGGAGACAAGAGTGAAACTCCATCTCAAAAAAAATAAACAA 97780

Qy        337 AAACAAAAAAA 347
              ||  ||| | |
Db      97779 AATAAAATACA 97769


RESULT 31
AC114480
LOCUS       AC114480               196773 bp    DNA     linear   PRI 29-JAN-2003
DEFINITION  Homo sapiens chromosome 3 clone RP11-755B10, complete sequence.
ACCESSION   AC114480
VERSION     AC114480.2  GI:28014590
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 196773)
  AUTHORS   Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
            Saenphimmachak,C., Buckley,D., Kibukawa,M., Raymond,C. and
            Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 196773)
  AUTHORS   Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2002) Genome Center, University of Washington,
            Box 352145, Seattle, WA 98195, USA
REFERENCE   3  (bases 1 to 196773)
  AUTHORS   Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
            Saenphimmachak,C., Buckley,D., Kibukawa,M., Raymond,C. and
            Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JAN-2003) Genome Center, University of Washington,
            Box 352145, Seattle, WA 98195, USA
COMMENT     On Jan 29, 2003 this sequence version replaced gi:19310299.
            -------------- Genome Center
                Center: University of Washington Genome Center
                Center Code: UWGC
                Web site: http://www.genome.washington.edu
                Contact: uwgchtgs@u.washington.edu
            -------------- Project Information
                Center project name: chr-3
                Center clone name: RP11-755B10 (bc0671)
            ------------- Summary Statistics
                Sequencing vector: unknown; 3% of reads
                Sequencing vector: plasmid; 97% of reads
                Chemistry: Dye-terminator ET; 94% of reads
                Chemistry: Dye-terminator Big Dye; 6% of reads
                Assembly program: Phrap; version 0.990319
                Consensus quality: 196438 bases at least Q40
                Consensus quality: 196751 bases at least Q30
                Consensus quality: 196772 bases at least Q20
                Insert size: 196773; sum-of-contigs
```

Quality coverage: 8.4x in Q20 bases; sum-of-contigs

Overlapping Sequences:
5': RP11-229A12 (UWGC:bc0315) AC092418, 11308-bp overlap
3': RP11-680P23 (UWGC:bc0527) AC099558, 118740-bp overlap

Sequence Quality Assessment:
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:
This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

EcoRI		HindIII		BglII	
SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt
8696	8856	3845	3999	10997	10643
6	<800	6382	6490	2067	2051
4995	4998	512	<800	12782	12807
2246	2241	449	<800	13590	12807
4061	4055	11421	11078	13168	12807
2665	2694	3245	3293	3447	3526
519	<800	3752	3826	5493	5591
1442	1386	1035	1032	1619	1545
6438	6457	881	873	9870	9617
2522	2531	5764	5726	6707	6693
9191	9289	6793	6788	3014	3167
2510	2531	889	873	4466	4431
138	<800	740	<800	16871	17658
2876	2910	2548	2558	1067	1017
1971	1982	2797	2775	5594	5591
812	796	5345	5298	10765	10643
1196	1166	552	<800	6602	6693
3348	3354	540	<800	4314	4431
11391	11222	1886	1862	846	836
16746	16576	5371	5298	3702	3743
340	<800	12306	11906	2801	2913
2626	2694	2938	2906	266	<800
4361	4390	321	<800	5324	5342
437	<800	14621	14206	3164	3280
2459	2531	7088	7077	7697	7733
18703	18633	684	<800	2197	2308
2240	2241	4160	4387	1501	1449
14113	13849	657	<800	1359	1311
27	<800	677	<800	15679	16033
1993	1982	669	<800	5370	5342
28749	28884	4437	4387	10381	10125
1970	1982	3581	3651	12752	12807
15338	15091	1088	1032		
417	<800	6707	6788		
21440	21694	6490	6490		
328	<800	4499	4387		
4697	4783	2212	2210		
478	<800	17962	17884		
987	967	4016	4136		
		2063	2063		
		8874	8921		
		3328	3518		
		22128	22660		
		7472	7554		
		1747	1718		

FEATURES Location/Qualifiers
source 1. .196773
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-755B10"
/clone_lib="RPCI human BAC library 11"

ORIGIN

Query Match 39.2%; Score 156.6; DB 5; Length 196773;
Best Local Similarity 71.4%; Pred. No. 2.1e-42;
Matches 222; Conservative 1; Mismatches 80; Indels 8; Gaps 1;

Qy 37 AAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAG 96

```
                ||| ||  |||  | ||    | ||  |    |      |    |||| |||||||| ||
Db      31090 AAAAATATATACAAAATACAGAAGCTGGCTGGGTGTGATGGCTCATGCTTGTAATCCTAG 31149

Qy         97 CACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGC 156
               ||||| ||||||||||||  || ||||||||||||||| ||| ||||||| | ||||||
Db      31150 CACTTTGGGAGGCCAAGGCAGGTGGATCACCTGAGGTCAGGAGTTCGAGACTAGCCTGGC 31209

Qy        157 CAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACA 216
              |||||||||||||||||||| |||||||||||||| | ||||||||| |||||||
Db      31210 CAACATGGTGAAACCCCGTCTCTACTAAAAATACAAACATTAGCTGGGCTTGGTGGCGGG 31269

Qy        217 CACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGA 276
              | | |||| |||||||||||| ||||| | ||| ||||||||||| |||||         |
Db      31270 CGCTTGTAATCCCAGCTACTGAGGAGGCAGAGGTTGCAGTGAGCCTAGATC--------A 31321

Qy        277 GCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAA 336
                |  |||       | | | ||  ||||||| ||:|| |||||||| || ||  ||
Db      31322 TGCCATTGCACTCTGGCCTGGGAGACAAGAGTGAAACTCCATCTCAAAAAAAATAAACAA 31381

Qy        337 AAACAAAAAAA 347
              ||  ||| | |
Db      31382 AATAAAATACA 31392


RESULT 32
AC016552/c
LOCUS       AC016552               102008 bp    DNA     linear   HTG 19-APR-2001
DEFINITION  Homo sapiens chromosome 5 clone CTC-285M15, WORKING DRAFT SEQUENCE,
            4 ordered pieces.
ACCESSION   AC016552
VERSION     AC016552.5  GI:7711562
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 102008)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 102008)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On May 6, 2000 this sequence version replaced gi:7710241.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 294922
            Center clone name: CIT-HSPC_285M15
            -------
            Summary Statistics
            Consensus quality: 92975 bases at least Q40
            Consensus quality: 98251 bases at least Q30
            Consensus quality: 100181 bases at least Q20
            Estimated insert size: 106000; pulse field gel estimation
            Estimated insert size: 101858; sum-of-contigs estimation
            Quality coverage:  5.46 in Q20 bases; pulse field gel estimation
            Quality coverage:  5.69 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    63901: contig of 63901 bp in length
            *    63902    64001: gap of unknown length
            *    64002    92415: contig of 28414 bp in length
            *    92416    92515: gap of unknown length
            *    92516   101204: contig of 8689 bp in length
            *   101205   101304: gap of unknown length
            *   101305   102008: contig of 704 bp in length.
FEATURES             Location/Qualifiers
     source          1. .102008
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTC-285M15"
                     /clone_lib="CalTech human BAC library C"
     gap             63902. .64001
                     /estimated_length=unknown
     gap             92416. .92515
                     /estimated_length=unknown
     gap             101205. .101304
                     /estimated_length=unknown
ORIGIN

  Query Match             39.1%;  Score 156.4;  DB 12;  Length 102008;
  Best Local Similarity   76.6%;  Pred. No. 1.8e-42;
  Matches  216;  Conservative    1;  Mismatches   62;  Indels    3;  Gaps    2;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||| |||| |||||||||||||||| ||||||||||||||| |||
Db      71686 ATGCCTGTAATCCCAGAACTTTGGGAGGCCAAGGTGGGTGGATCACCTGAGGTCAGGAGT 71627

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||||||||||| |||||||||| ||| ||||||||||||||||| ||||
Db      71626 TCGAGACCAGCCTGGCCAACATAGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 71567

Qy        201 TGGGCATGGT--GGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGA 258
                ||  |||| ||||||| |||||||||||||||||||| |||| | |||   ||||
Db      71566 CAGGTGTGGTGGGGCACACGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGCAGAAT 71507

Qy        259 GCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGT 318
                || | |||   | | |    |||      | | |  | ||||||||||||:||||
Db      71506 CGCTTGAACCCAGGAGGTGGAGCCTGCACTCCAGCCTG-GGCAACAGAGTGAGACTCCGT 71448

Qy        319 CTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 360
              ||||||||  | ||||||| | ||||||   |||||   |||
Db      71447 CTCAAAAAACAAAACAAAACAAAAAAAAGTAATAAGTGCTTG 71406


RESULT 33
AC008499/c
LOCUS       AC008499               244525 bp    DNA     linear   PRI 26-JAN-2002
DEFINITION  Homo sapiens chromosome 5 clone CTC-43803, complete sequence.
ACCESSION   AC008499
VERSION     AC008499.8  GI:18376856
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 244525)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 244525)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 244525)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-DEC-2001) DOE Joint Genome Institute, 2800 Mitchell
```

```
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 244525)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jan 26, 2002 this sequence version replaced gi:17386226.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.7% of Sequence;
            Estimated Total Number of Errors is 0.4.
FEATURES             Location/Qualifiers
     source          1. .244525
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTC-438O3"
ORIGIN

  Query Match             39.1%;  Score 156.4;  DB 5;  Length 244525;
  Best Local Similarity   76.6%;  Pred. No. 2.7e-42;
  Matches  216;  Conservative    1;  Mismatches   62;  Indels    3;  Gaps    2;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||| |||| |||||||||||||||| |||||||||||||||| |||
Db       86801 ATGCCTGTAATCCCAGAACTTTGGGAGGCCAAGGTGGGTGGATCACCTGAGGTCAGGAGT 86742

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               ||||||||| |||||||||||| |||||||||| ||| ||||||||||||||||| ||||
Db       86741 TCGAGACCAGCCTGGCCAACATAGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGCC 86682

Qy         201 TGGGCATGGT--GGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGA 258
                 ||  ||||  ||||||| |||||||||||||||||||| |||| | |||   ||||
Db       86681 CAGGTGTGGTGGGGCACACGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGCAGAAT 86622

Qy         259 GCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGT 318
                 || | |||   | |  |    |||      | | |  | ||||||||||:||||
Db       86621 CGCTTGAACCCAGGAGGTGGAGCCTGCACTCCAGCCTG-GGCAACAGAGTGAGACTCCGT 86563

Qy         319 CTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 360
               |||||||  | |||||| | ||||||   |||||   |||
Db       86562 CTCAAAAAACAAAACAAAACAAAAAAAAGTAATAAGTGCTTG 86521


RESULT 34
AC107377
LOCUS       AC107377               88848 bp    DNA     linear   PRI 01-MAY-2002
DEFINITION  Homo sapiens chromosome 17, clone CTD-2314M10, complete sequence.
ACCESSION   AC107377
VERSION     AC107377.4  GI:20377031
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 88848)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone CTD-2314M10
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 88848)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JAN-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 88848)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 88848)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On May 1, 2002 this sequence version replaced gi:20303884.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ---------- Genome Center
```

```
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
          -------------- Project Information
            Center project name: L21459
            Center clone name: 2314_M_10
          --------------.
FEATURES             Location/Qualifiers
     source          1. .88848
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="CTD-2314M10"
                     /clone_lib="CITD1 Human BAC"
     repeat_region   complement(98. .394)
                     /rpt_family="AluSc"
     repeat_region   408. .446
                     /rpt_family="L2"
     repeat_region   complement(1891. .2157)
                     /rpt_family="AluSx"
     repeat_region   2253. .2538
                     /rpt_family="L2"
     repeat_region   complement(2618. .2816)
                     /rpt_family="AluSq"
     repeat_region   2882. .3044
                     /rpt_family="L2"
     repeat_region   3053. .3253
                     /rpt_family="MIR"
     repeat_region   complement(4123. .4358)
                     /rpt_family="AluSx"
     repeat_region   complement(4539. .4848)
                     /rpt_family="AluJb"
     repeat_region   complement(5101. .5396)
                     /rpt_family="AluSc"
     repeat_region   complement(5449. .5754)
                     /rpt_family="AluSq"
     repeat_region   complement(5927. .6211)
                     /rpt_family="AluSq"
     repeat_region   complement(6398. .6702)
                     /rpt_family="AluJo"
     repeat_region   complement(6715. .7020)
                     /rpt_family="AluSq"
     repeat_region   complement(7275. .7572)
                     /rpt_family="AluSx"
     repeat_region   7725. .7745
                     /rpt_family="AT_rich"
     repeat_region   8488. .8513
                     /rpt_family="(TTTA)n"
     repeat_region   complement(8517. .8666)
                     /rpt_family="FRAM"
     repeat_region   8667. .8695
                     /rpt_family="(TTCA)n"
     repeat_region   complement(8714. .9005)
                     /rpt_family="AluSc"
     repeat_region   complement(9022. .9328)
                     /rpt_family="AluSx"
     repeat_region   9492. .9785
                     /rpt_family="AluSq"
     repeat_region   complement(9849. .10156)
                     /rpt_family="AluJb"
     repeat_region   11807. .12100
                     /rpt_family="AluSx"
     repeat_region   complement(12718. .12832)
                     /rpt_family="MIR"
     repeat_region   13094. .13244
                     /rpt_family="FRAM"
     repeat_region   13367. .13673
                     /rpt_family="AluY"
     repeat_region   13686. .13997
                     /rpt_family="AluJo"
     repeat_region   14455. .14475
                     /rpt_family="AT_rich"
     repeat_region   complement(14478. .14608)
                     /rpt_family="FLAM_C"
     repeat_region   14917. .14938
                     /rpt_family="(A)n"
     repeat_region   14946. .15215
                     /rpt_family="AluJb"
     repeat_region   15576. .15735
                     /rpt_family="AluSx"
     repeat_region   15736. .15982
                     /rpt_family="AluSx"
     repeat_region   16211. .16518
                     /rpt_family="AluSx"
     unsure          16519. .16524
                     /note="<30 Qual SNGL Region"
     repeat_region   16635. .16933
                     /rpt_family="AluJb"
     repeat_region   complement(17254. .17410)
                     /rpt_family="MER58A"
     repeat_region   17422. .17547
                     /rpt_family="L2"
     repeat_region   complement(17944. .18240)
                     /rpt_family="AluSx"
     repeat_region   complement(18426. .18746)
                     /rpt_family="AluSx"
     repeat_region   complement(18751. .19065)
                     /rpt_family="AluSx"
     repeat_region   19095. .19142
                     /rpt_family="AT_rich"

  Query Match             39.0%;  Score 155.8;  DB 5;  Length 88848;
  Best Local Similarity   74.9%;  Pred. No. 2.7e-42;
  Matches  212;  Conservative    0;  Mismatches   62;  Indels    9;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||| | | ||||| ||||||| | ||||||||||||||||||||||| |||
Db      57874 ACGCCTGTAATTCTACCACTTTGGGAGGCTGAAGTGGGCGGATCACCTGAGGTCAGGAGT 57933

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||||||||||||| |||| || |||| |||||||||||||||| ||||
Db      57934 TCGAGACCAGCCTGGCCAACATGGCGAAATCCTGTCTCTACTAAAAATACAAAAATTAGC 57993

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||  |||| |    | |||||| |||||||||||||||| ||||| ||||||||||||
Db      57994 TGGGTGTGGTTGTGGGCGCCTGTAATCCCAGCTACTCAGGAGGCGGAGGTTGCAGTGAGC 58053

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||||| |          |   | |   | | |         |||||  |||| | ||||
Db      58054 TGAGATTGTGCCACTGCACTCCAGCCTGGGTGA---------CAGAGCAAGACCCTGTCT 58104

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCC 363
              |||||| || || |||||| |||||||| | ||  ||   |||
Db      58105 CAAAAAAAAAAAAAAAAAAAAAAAAAAACAAAAACCCAAAATCC 58147


RESULT 35
AC172796/c
LOCUS       AC172796              182725 bp    DNA     linear   HTG 31-JAN-2006
DEFINITION  Callicebus moloch clone LB5-41K3, WORKING DRAFT SEQUENCE, 9 ordered
            pieces.
ACCESSION   AC172796
VERSION     AC172796.2  GI:86203525
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Callicebus moloch (Dusky titi)
  ORGANISM  Callicebus moloch
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Platyrrhini;
            Cebidae; Callicebinae; Callicebus.
REFERENCE   1  (bases 1 to 182725)
  AUTHORS   Antonellis,A., Bass,D., Benjamin,B., Bera,J., Blakesley,R.W.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H.,
            Franks,S., Fuksenko,T., Gestole,M., Greene,A., Guan,X., Gupta,J.,
```

```
          Gurson,N., Haghighi,P., Han,E., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
          Hunter,G., Hurle,B., Idol,J.R., Kwong,P., Laric,P., Larson,S.,
          Lee-Lin,S.-Q., Legaspi,R., Madden,M., Maduro,Q.L., Maduro,V.B.,
          Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
          Montemayor,C., Mullikin,J.C., Park,M., Prasad,A., Puri,O.,
          Rantz,K., Reddix-Dugue,N., Sante,A., Schandler,K., Schueler,M.G.,
          Sison,C., Stantripop,S., Taye,A., Thomas,J.W., Thomas,P.J.,
          Tsipouri,V., Ung,L., Vogt,J.L., Wetherby,K.D., Withers,T.R.,
          Young,A. and Green,E.D.
  TITLE   NISC Comparative Sequencing Initiative
  JOURNAL Unpublished
REFERENCE 2  (bases 1 to 182725)
  AUTHORS Green,E.D.
  TITLE   Direct Submission
  JOURNAL Submitted (23-NOV-2005) NIH Intramural Sequencing Center, 5625
          Fishers Lane, Rockville, MD 20852, USA
REFERENCE 3  (bases 1 to 182725)
  AUTHORS Green,E.D.
  TITLE   Direct Submission
  JOURNAL Submitted (31-JAN-2006) NIH Intramural Sequencing Center, 5625
          Fishers Lane, Rockville, MD 20852, USA
COMMENT   On Jan 31, 2006 this sequence version replaced gi:82617757.
          -------------- Genome Center
               Center: NIH Intramural Sequencing Center
               Center code: NISC
               Web site: http://www.nisc.nih.gov
               Contact: nisc_zoo@nhgri.nih.gov
          -------------- Project Information
               Center project name: mrn
               Center clone name: 041K03

          The sequence data in this record represents an 'enhanced'
          version of a Phase 2 submission. Specifically, the indicated
          order and orientation of each sequence contig has been
          established using one or more of the following: read-pair
          data from individual subclones, overlaps with neighboring
          clones, alignment with available reference sequence (e.g.,
          human), and/or confirmation by PCR testing. In addition,
          the sequence assembly is generally based on at least 8X average
          coverage in Q20 bases and has been reviewed to rule out
          gross misassemblies, the low-quality ends of sequence
          contigs have been trimmed away, and each base is associated
          with a Phrap-derived quality score.
          -------------- Summary Statistics
               Sequencing vector: plasmid; n/a; 100% of reads
               Chemistry: Dye-terminator Big Dye; 100% of reads
               Assembly program: Phrap; version 0.990319
               Consensus quality: 179767 bases at least Q40
               Consensus quality: 180990 bases at least Q30
               Consensus quality: 181586 bases at least Q20
               Insert size: 214000; agarose-fp
               Insert size: 181925; sum-of-contigs
               Quality coverage: 9.44x in Q20 bases; agarose-fp
               Quality coverage: 11.11x in Q20 bases; sum-of-contigs
          --------------.
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 9 contigs. Gaps between the contigs
          * are represented as runs of N. The order of the pieces
          * is believed to be correct as given, however the sizes
          * of the gaps between them are based on estimates that have
          * provided by the submittor.
          * This sequence will be replaced
          * by the finished sequence as soon as it is available and
          * the accession number will be preserved.
          *        1    22543: contig of 22543 bp in length
          *    22544    22643: gap of unknown length
          *    22644    31875: contig of 9232 bp in length
          *    31876    31975: gap of unknown length
          *    31976   122998: contig of 91023 bp in length
          *   122999   123098: gap of unknown length
          *   123099   136487: contig of 13389 bp in length
          *   136488   136587: gap of unknown length
          *   136588   139369: contig of 2782 bp in length
          *   139370   139469: gap of unknown length
          *   139470   148608: contig of 9139 bp in length
          *   148609   148708: gap of unknown length
          *   148709   157275: contig of 8567 bp in length
          *   157276   157375: gap of unknown length
          *   157376   179801: contig of 22426 bp in length
          *   179802   179901: gap of unknown length
          *   179902   182725: contig of 2824 bp in length.
FEATURES             Location/Qualifiers
     source          1. .182725
                     /organism="Callicebus moloch"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9523"
                     /clone="LB5-41K3"
                     /clone_lib="LB5"
                     /note="BAC resource: http://bacpac.chori.org/"
     misc_feature    1. .22543
                     /note="assembly_fragment
                     missing T7 clone end on 5' end of insert"
     misc_feature    1. .18769
                     /note="clone overlaps with GenBank Accession Number
                     AC174418 clone LB5-285I22 (center project name mrm)"
     gap             22544. .22643
                     /estimated_length=unknown
     misc_feature    22644. .31875
                     /note="assembly_fragment"
     gap             31876. .31975
                     /estimated_length=unknown
     misc_feature    31976. .122998
                     /note="assembly_fragment"
     misc_feature    50743. .182725
                     /note="clone overlaps with GenBank Accession Number
                     AC172795 clone LB5-290P9 (center project name mro)"
     gap             122999. .123098
                     /estimated_length=unknown
     misc_feature    123099. .136487
                     /note="assembly_fragment"
     gap             136488. .136587
                     /estimated_length=unknown
     misc_feature    136588. .139369
                     /note="assembly_fragment"
     gap             139370. .139469
                     /estimated_length=unknown
     misc_feature    139470. .148608
                     /note="assembly_fragment"
     gap             148609. .148708
                     /estimated_length=unknown
     misc_feature    148709. .157275
                     /note="assembly_fragment"
     gap             157276. .157375
                     /estimated_length=unknown
     misc_feature    157376. .179801
                     /note="assembly_fragment"
     gap             179802. .179901
                     /estimated_length=unknown
     misc_feature    179902. .182725
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
ORIGIN

  Query Match            39.0%;  Score 155.8;  DB 12;  Length 182725;
  Best Local Similarity  70.5%;  Pred. No. 3.9e-42;
  Matches  225;  Conservative    1;  Mismatches   83;  Indels   10;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| ||||||| || |||| | |||||||||||||  ||
Db      23508 ACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGATGGGTGAATCACCTGAGGTCAGAAGT 23449

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||| | |||||||||||||||||| ||||| |||||||||||||  || ||||
Db      23448 TAGAGACCAGCTTGGCCAACATGGTGAAACCTCGTCTCTACTAAAAATACAGGAATTAGC 23389
```

```
Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||||||||||||| ||||| ||||||||||||||  |||| ||||| || ||||||||
Db      23388 TGGGCATGGTGGCACGCACCTATAGTCCCAGCTACTTGGGAGGCGGAGGTTACAGTGAGC 23329

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              ||||| ||          ||   ||   ||    | | ||||| |||||: | |||
Db      23328 CGAGATCACA----------CAACTGCACTCCAGCCTGGCAACAGAGGGAGACTTCATCT 23279

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGAC 380
              ||||||||| ||  ||| |  | ||      | ||       |  |||   | ||   ||
Db      23278 CAAAAACAAAAAACAAAGAGTATAACTGGGCATGATGGCTCACGCCTGTAATCCCATCAC 23219

Qy        381 TATTGCAGGAGACCAAAAA 399
              | | | |||  |   | ||
Db      23218 TTTGGAAGGCCAAGGAGAA 23200
```

RESULT 36
DJ534K4
LOCUS DJ534K4 216387 bp DNA linear PRI 23-DEC-1998
DEFINITION Homo sapiens S164 gene, partial cds; PS1 and hypothetical protein genes, complete cds; and S171 gene, partial cds.
ACCESSION AF109907
VERSION AF109907.1 GI:4050085
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 216387)
AUTHORS Rowen,L., Madan,A., Qin,S., Abbasi,N., Dors,M., Ratcliffe,A., Madan,A., Dickhoff,R., Shaffer,T., James,R., Lasky,S. and Hood,L.
TITLE Complete sequence of the gene for presenilin 1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 216387)
AUTHORS Rowen,L.
TITLE Direct Submission
JOURNAL Submitted (30-NOV-1998) Department of Molecular Biotechnology, Box 357730 University of Washington, Seattle, Washington 98195, USA
COMMENT Sequencing methodology: high redundancy shotgun using plasmids. Interspersed Repeats were identified with RepeatMasker (available from http://ftp.genome.washington.edu/RM/RepeatMasker.html).
FEATURES Location/Qualifiers

```
     source          1. .216387
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /map="14q24.3"
                     /clone="BAC534K4"
                     /clone_lib="P. deJong RPC11"
     repeat_region   complement(262. .556)
                     /rpt_family="AluJo"
     repeat_region   750. .1064
                     /rpt_family="AluSg"
     mRNA            join(<1465. .1522,5979. .6139,14890. .15075,17363. .17416,
                     17650. .17733,21177. .21463,23822. .24044,24159. .24248,
                     25789. .25862,27305. .27455,28794. .29118,29491. .29765,
                     30115. .30213,32250. .32297,37674. .>37766)
                     /product="S164"
     CDS             join(<1465. .1522,5979. .6139,14890. .15075,17363. .17416,
                     17650. .17733,21177. .21463,23822. .24044,24159. .24248,
                     25789. .25862,27305. .27455,28794. .29118,29491. .29765,
                     30115. .30213,32250. .32297,37674. .37766)
                     /note="unknown; intron-exon boundaries defined by ESTs
                     AI247672, AA677603, AA677761, T40389, AA546306, AA147425,
                     AA147323, AA910441, AA170655, and cDNA in L40392; the
                     sequence in L40392 begins at 28812; the exons from
                     21177-21463 and 23822-24044 contain a repeat; only 1 EST,
                     T40389 matches the exon from 21177-21463; since several 3'
                     ESTs end in the beginning of this exon, and several 5'
                     ESTs start at the end of this exon, it is possible that
                     two genes are in this span; if so, the stop codon for the
                     first gene could not be identified; the beginning and the
                     end of this gene match a hypothetical 'RNA recognition
                     protein' from several species, based on blastx"
                     /codon_start=1
                     /product="S164"
                     /protein_id="AAC97961.1"
                     /db_xref="GI:4050087"
                     /translation="KCGLVLSWKRVQGASGKLQAFGFCEYKEPESTLRALRLLHDLQI
                     GEKKLLVKVDAKTKAQLDEWKAKKKASNGNARPETVTNDDEEALDEETKRRDQMIKGA
                     IEVLIREYSSELNAPSQESDSHPRKKKKEKKEDIFRRFPVAPLIPYPLITKEDINAIE
                     MEEDKRDLISREISKFRDTHKKLEEEKGKKEKERQEIEKERRERERERERERERERE
                     RERERERERERERERERERERERERERERERERERERERERERERERERERERERERE
                     REKSRDRERERERERERERERERERERERERERERERERERERERERERERERERERE
                     RRKLERKLREKEAAYQERLKNWEIRERKKTREYEKEAEREEERRREMAKEAKRLKEFL
                     EDYDDDRDDPKYYRGSALQKRLRDREKEMEADERDRKREKEELEEIRQRLLAEGHPDP
                     DAELQRMEQEAERRRQPQIKQEPESEEEEEEKQEKEEKREEPMEEEEEPEQKPCLKPT
                     LRPISSAPSVSSASGNATPNTPGDESPCGIIIPHENSPDQQQPEEHRPKIGLSLKLGA
                     SNSPGQPNSVKRKKLPVDSVFNKFEDEDSDDVPRKRKLVPLDYGEDDKNATKGTVNTE
                     EKRKHIKSLIEKIPTAKPELFAYPLDWSIVDSILMERRIRPWINKKIIEYIGEEEATL
                     VDFVCSKVMAHSSPQSILDDVAMVLDEEAEVFIVKMWRLLIYETEAKKIGLVK"
     repeat_region   complement(1870. .2168)
                     /rpt_family="AluSx"
     repeat_region   complement(2523. .2820)
                     /rpt_family="AluSp"
     repeat_region   complement(2992. .3287)
                     /rpt_family="AluSc"
     repeat_region   complement(3417. .3716)
                     /rpt_family="AluSq"
     repeat_region   complement(3956. .4121)
                     /rpt_family="L1MC4"
     repeat_region   4141. .4426
                     /rpt_family="AluSc"
     repeat_region   complement(4676. .4975)
                     /rpt_family="AluSx"
     repeat_region   complement(5052. .5163)
                     /rpt_unit_seq="taaaa"
     repeat_region   complement(5372. .5672)
                     /rpt_family="AluY"
     repeat_region   6838. .7145
                     /rpt_family="AluSq"
     repeat_region   7176. .7244
                     /rpt_unit_seq="taaaa"
     repeat_region   complement(7334. .7632)
                     /rpt_family="AluSx"
     repeat_region   complement(7643. .7925)
                     /rpt_family="AluSq"
     repeat_region   complement(8038. .8328)
                     /rpt_family="AluSq"
     repeat_region   complement(8864. .8890)
                     /rpt_family="AT_rich"
     repeat_region   complement(9200. .9490)
                     /rpt_family="AluJb"
     repeat_region   9570. .9864
                     /rpt_family="AluSx"
     repeat_region   9894. .9921
                     /rpt_family="AT_rich"
     repeat_region   10178. .10277
                     /rpt_unit_seq="ggga"
     repeat_region   complement(10572. .10696)
                     /rpt_family="L1PA4"
     repeat_region   complement(10714. .10844)
                     /rpt_family="FLAM_C"
     repeat_region   complement(11506. .11784)
                     /rpt_family="AluSq"
     repeat_region   complement(11928. .12230)
                     /rpt_family="AluY"
     repeat_region   complement(12429. .12653)
                     /rpt_family="AluY"
     unsure          12447. .12450
                     /note="sequence could be aacgg--possible compression"
     repeat_region   complement(12658. .12788)
                     /rpt_family="AluSg"
```

```
repeat_region   complement(12789. .12839)
                /rpt_unit_seq="ca"
repeat_region   complement(12849. .13126)
                /rpt_family="AluSg"
repeat_region   13302. .13597
                /rpt_family="AluY"
repeat_region   complement(13960. .14251)
                /rpt_family="AluSg"
repeat_region   complement(15155. .15181)
                /rpt_family="POLY_A"
repeat_region   complement(15182. .15481)
                /rpt_family="AluY"
repeat_region   complement(15484. .15617)
                /rpt_family="AluJb"
repeat_region   15966. .16094
                /rpt_family="FLAM_C"
repeat_region   complement(16115. .16401)
                /rpt_family="AluJb"
repeat_region   complement(16498. .16518)
                /rpt_family="AT_rich"
repeat_region   complement(16645. .16869)
                /rpt_family="AluY"
repeat_region   complement(17813. .17846)
                /rpt_family="AT_rich"
repeat_region   19444. .19528
                /rpt_unit_seq="ta"
repeat_region   complement(20193. .20244)
                /rpt_unit_seq="caaaa"
repeat_region   complement(20503. .20523)
                /rpt_family="AT_rich"
repeat_region   21182. .21305
                /rpt_unit_seq="gaaa"
repeat_region   21225. .21346
                /rpt_unit_seq="ga"
repeat_region   complement(21557. .21838)
                /rpt_family="AluSx"
STS             21720. .22035
                /standard_name="D14S77"
repeat_region   complement(21865. .21981)
                /rpt_unit_seq="ga"
repeat_region   22297. .22330
                /rpt_family="AT_rich"
repeat_region   complement(22333. .22627)
                /rpt_family="AluSx"
repeat_region   22982. .23028
                /rpt_unit_seq="ca"
repeat_region   complement(23090. .23230)
                /rpt_family="FRAM"
repeat_region   23832. .23950
                /rpt_unit_seq="ga"
repeat_region   23905. .24029
                /rpt_unit_seq="gaaa"
repeat_region   24180. .24254
                /rpt_unit_seq="gaaaa"
repeat_region   complement(24756. .24802)
                /rpt_unit_seq="ta"
repeat_region   complement(24804. .25103)
                /rpt_family="AluSx"
repeat_region   complement(26155. .26309)
                /rpt_family="AluJo"
repeat_region   27475. .27742
                /rpt_family="AluY"
repeat_region   complement(27752. .27844)
                /rpt_family="AluSx"
repeat_region   complement(27846. .28147)
                /rpt_family="AluSx"
repeat_region   complement(28148. .28360)
                /rpt_family="AluJb"
repeat_region   complement(28394. .28691)
                /rpt_family="AluSx"
repeat_region   28837. .28930
                /rpt_unit_seq="gaa"
repeat_region   complement(30243. .30267)
```

```
  Query Match            39.0%;  Score 155.8;  DB 5;  Length 216387;
  Best Local Similarity  77.0%;  Pred. No. 4.2e-42;
  Matches  218;  Conservative    1;  Mismatches   53;  Indels   11;  Gaps    2;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||| |||||||| ||||||||||||||||||||
Db     195538 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACCTGAGGTCAAGAGA 195597

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||| |||||||||||||||||||||||||||||||| |||||||||||||||| ||||
Db     195598 TCGAGACTATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 195657

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-------GCCGGAGATTGC 253
              ||||| |||||| |  || |||||||||||||||||| |||       ||  |||| |
Db     195658 TGGGCGTGGTGGTATGCATCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAAGAGAGTTG 195717

Qy        254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGAT----CACAGAGTGAGCAGAGTG 309
                ||| |      |||||| | || |  | |  ||||    ||          || ||| |
Db     195718 CTTGAACCTGAGAGGCAGAGGGTGCAGTGAACCAAGATACTCCAGCCTAGTGACACAGCG 195777

Qy        310 AGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATA 352
              ||||:|||||||||||| || || |||||| ||||||  | ||
Db     195778 AGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAACACCTA 195820


RESULT 37
AL732374/c
LOCUS       AL732374              224187 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP13-444K19 on chromosome X Contains
            a mitochondrial ribosomal protein S18C (MRPS18C) pseudogene, the 3'
            end of the gene for a novel protein similar to PHD finger protein 2
            PHF2 and a CpG island, complete sequence.
ACCESSION   AL732374
VERSION     AL732374.14  GI:23476649
KEYWORDS    HTG; CpG island; MRPS18C; PHF2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 224187)
  AUTHORS   Chapman,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 2, 2002 this sequence version replaced gi:23393869.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome X, constructed by the Sanger Centre Chromosome X Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/ChrX
            RP13-444K19 is from the library RPCI-13.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            -------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
```

```
        except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .224187
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP13-444K19"
                     /clone_lib="RPCI-13.2"
     gene            11617. .12039
                     /locus_tag="RP13-444K19.1-001"
                     /pseudo
     CDS             11617. .12039
                     /locus_tag="RP13-444K19.1-001"
                     /note="match: proteins: Q8R2L5 Q9Y3D5"
                     /pseudo
                     /codon_start=1
                     /product="mitochondrial ribosomal protein S18C (MRPS18C)
                     pseudogene"
     polyA_site      complement(122345)
     gene            join(complement(AL589872.9:883. .971),
                     complement(Z98051.7:6109. .6298),
                     complement(208423. .208508),complement(207924. .208032),
                     complement(203326. .203486),complement(202260. .202401),
                     complement(200042. .200228),complement(196787. .196949),
                     complement(188260. .188347),complement(187820. .187926),
                     complement(185535. .185626),complement(181358. .181447),
                     complement(179266. .179568),complement(178401. .178504),
                     complement(173431. .173609),complement(172743. .172828),
                     complement(171481. .171614),complement(170579. .170892),
                     complement(148509. .148604),complement(129799. .129908),
                     complement(125845. .126181),complement(122347. .124911))
                     /gene="PHF8"
                     /locus_tag="RP13-444K19.2-001"
     mRNA            join(complement(AL589872.9:883. .971),
                     complement(Z98051.7:6109. .6298),
                     complement(208423. .208508),complement(207924. .208032),
                     complement(203326. .203486),complement(202260. .202401),
                     complement(200042. .200228),complement(196787. .196949),
                     complement(188260. .188347),complement(187820. .187926),
                     complement(185535. .185626),complement(181358. .181447),
                     complement(179266. .179568),complement(178401. .178504),
                     complement(173431. .173609),complement(172743. .172828),
                     complement(171481. .171614),complement(170579. .170892),
                     complement(148509. .148604),complement(129799. .129908),
                     complement(125845. .126181),complement(122347. .124911))
                     /gene="PHF8"
                     /locus_tag="RP13-444K19.2-001"
                     /product="PHD finger protein 8"
                     /note="match: ESTs: AA551120.1 AL707548.1 AU118751.1
                     AW249934.1 AW304438.1 BG719138.1 BG719444.1 BG939390.1
                     BI333269.1 BI829064.1 BM462299.1 BQ025306.1 BX091300.1
                     CA314327.1
                     match: cDNAs: AB029034.1 AK022788.1"
     gene            complement(join(123382. .123724,124827. .124911,
                     125845. .126181,129799. .129908,148509. .148604,
                     170579. .170892,171481. .171614,172743. .172828,
                     173431. .173609,178401. .178504,179266. .179568,
                     181358. .181447,185535. .185626,187820. .187926,
                     188260. .188347,196787. .196949,200042. .200228,
                     202260. .202401,203326. .203486,207924. .>207927))
                     /gene="PHF8"
                     /locus_tag="RP13-444K19.2-007"
     mRNA            complement(join(123382. .123724,124827. .124911,
                     125845. .126181,129799. .129908,148509. .148604,
                     170579. .170892,171481. .171614,172743. .172828,
                     173431. .173609,178401. .178504,179266. .179568,
                     181358. .181447,185535. .185626,187820. .187926,
                     188260. .188347,196787. .196949,200042. .200228,
                     202260. .202401,203326. .203486,207924. .>207927))
                     /gene="PHF8"
                     /locus_tag="RP13-444K19.2-007"
                     /product="PHD finger protein 8"
                     /note="match: ESTs: AA620622.1 AW386032.1 AW934879.1
                     BF772185.1 BF772623.1 BG195947.1 BI828904.1
                     match: cDNAs: BC017720.1"
     CDS             complement(join(123646. .123724,124827. .124911,
                     125845. .126181,129799. .129908,148509. .148604,
                     170579. .170892,171481. .171614,172743. .172828,
                     173431. .173609,178401. .178504,179266. .179568,
                     181358. .181447,185535. .185626,187820. .187926,
                     188260. .188347,196787. .196949,200042. .200228,
                     202260. .202401,203326. .203486,207924. .>207927))
                     /gene="PHF8"
                     /locus_tag="RP13-444K19.2-007"
                     /standard_name="OTTHUMP00000023367"
                     /codon_start=3
                     /product="PHD finger protein 8"
                     /protein_id="CAI41582.1"
                     /db_xref="GI:57209883"
                     /db_xref="UniProtKB/TrEMBL:Q5JPS2"
                     /translation="SSDEVILKPTGNQLTVEPLEENSFSVPILVLKKDGLGMTLPSPS
                     FTVRDVEHYVGSDKEIDVIDVTRQADCKMKLGDFVKYYYSGKREKVLNVISLEFSDTR
                     LSNLVETPKIVRKLSWVENLWPEECVFERPNVQKYCLMSVRDSYTDFHIDFGGTSVWY
                     HVLKGEKIFYLIRPTNANLTLFECWSSSSNQNEMFFGDQVDKCYKCSVKQGQTLFIPT
                     GWIHAVLTPVDCLAFGGNFLHSLNIEMQLKAYEIEKRLSTADLFRFPNFETICWYVGK
                     HILDIFRGLRENRRHPASYLVHGGKALNLAFRAWTRKEALPDHEDEIPETVRTVQLIK
                     DLAREIRLVEDIFQQNVGKTSNIFGLQRIFPAGSIPLTRPAHSTSVSMSRLSLPSKNG
                     SKKKGLKPKELFKKAERKGKESSALGPAGQLSYNLMDTYSHQALKTGSFQKAKFNITG
                     ACLNDSDDDSPDLDLDGNESPLALLMSNGSTKRVKSLSKSRRTKIAKKVDKARLMAEQ
                     VMEDEFDLDSDDELQIDERLGKEKATLIIRPKFPRKLPRAKPCSDPNRVREPGEVEFD
                     IEEDYTTDEDMVEGVEGKLGNGSGAGGILDLLKASRQVGGPDYAALTEAPASPSTQEA
                     IQGMLCMANLQSSSSSPATSSLQAWWTGGQDRSSGSSSSGLGTVSNSPASQRTPGKRP
                     IKRPAYWRTESEEEEENASLDEQDSLGACFKDAEYIYPSLESDDDDPALKSRPKKKKN
                     SDDAPWSPKARVTPTLPKQDRPVREGTRVASIETGLAAAAAKLAQQELQKAQKKKYIK
                     KKPLLKEVEQPRPQDSNLSLTVPAPTVAATPQLVTSSSPLPPPEPKQEALSGSLADHE
                     YTARPNAFGMAQANRSTTPMAPGVFLTQRRPSVGSQSNQAGQGKRPKKGLATAKQRLG
                     RILKIHRNGKLLLRQVIVQAECRQAIHEPKLKRRDAHP"
     CDS             join(complement(AL589872.9:883. .898),
                     complement(Z98051.7:6109. .6298),
                     complement(208423. .208508),complement(207924. .208032),
                     complement(203326. .203486),complement(202260. .202401),
                     complement(200042. .200228),complement(196787. .196949),
                     complement(188260. .188347),complement(187820. .187926),
                     complement(185535. .185626),complement(181358. .181447),
                     complement(179266. .179568),complement(178401. .178504),
                     complement(173431. .173609),complement(172743. .172828),
                     complement(171481. .171614),complement(170579. .170892),
                     complement(148509. .148604),complement(129799. .129908),
                     complement(125845. .126181),complement(124823. .124911))
                     /gene="PHF8"
                     /locus_tag="RP13-444K19.2-001"
                     /standard_name="OTTHUMP00000023362"
                     /note="match: proteins: BAC65729 Q9UPP1"
                     /codon_start=1
                     /product="PHD finger protein 8"
                     /protein_id="CAI41576.1"
                     /db_xref="GI:57209878"
                     /db_xref="GOA:Q5VUJ4"
                     /db_xref="InterPro:IPR001965"
                     /db_xref="InterPro:IPR003347"
                     /db_xref="InterPro:IPR006209"
                     /db_xref="UniProtKB/TrEMBL:Q5VUJ4"
                     /translation="MNRSRAIVQRGRVLPPPAPLDTTNLAGRRTLQGRAKMASVPVYC
                     LCRLPYDVTRFMIECDMCQDWFHGSCVGVEEEKAADIDLYHCPNCEVLHGPSIMKKRR
                     GSSKGHDTHKGKPVKTGSPTFVRELRSRTFDSSDEVILKPTGNQLTVEFLEENSFSVP
                     ILVLKKDGLGMTLPSPSFTVRDVEHYVGSDKEIDVIDVTRQADCKMKLGDFVKYYYSG
                     KREKVLNVISLEFSDTRLSNLVETPKIVRKLSWVENLWPEECVFERPNVQKYCLMSVR
                     DSYTDFHIDFGGTSVWYHVLKGEKIFYLIRPTNANLTLFECWSSSSNQNEMFFGDQVD
                     KCYKCSVKQGQTLFIPTGWIHAVLTPVDCLAFGGNFLHSLNIEMQLKAYEIEKRLSTA
                     DLFRFPNFETICWYVGKHILDIFRGLRENRRHPASYLVHGGKALNLAFRAWTRKEALP
                     DHEDEIPETVRTVQLIKDLAREIRLVEDIFQQNVGKTSNIFGLQRIFPAGSIPLTRPA
                     HSTSVSMSRLSLPSKNGSKKKGLKPKELFKKAERKGKESSALGPAGQLSYNLMDTYSH
                     QALKTGSFQKAKFNITGACLNDSDDDSPDLDLDGNESPLALLMSNGSTKRVKSLSKSR
                     RTKIAKKVDKARLMAEQVMEDEFDLDSDDELQIDERLGKEKATLIIRPKFPRKLPRAK
                     PCSDPNRVREPGEVEFDIEEDYTTDEDMVEGVEGKLGNGSGAGGILDLLKASRQVGGP
                     DYAALTEAPASPSTQEAIQGMLCMANLQSSSSSPATSSLQAWWTGGQDRSSGSSSSGL
```

```
                     GTVSNSPASQRTPGKRPIKRPAYWRTESEEEEENASLDEQDSLGACFKDAEYIYPSLE
                     SDDDDPALKSRPKKKKNSDDAPWSPKARVTPTLPKQDRPVREGTRVASIETGLAAAAA
                     KLAQQELQKAQKKKYIKKKPLLKEVEQPRPQDSNLSLTVPAPTVAATPQLVTSSSPLP
                     PPEPKQEALSGSLADHEYTARPNAFGMAQANRSTTPMAPGVFLTQRRPSVGSQSNQAG
                     QGKRPKKGLATAKQRLGRILKIHRNGKLLL"
     gene            complement(join(128280. .129067,129799. .129908,
                     148509. .148604,170579. .170892,171481. .171614,
                     172743. .172828,173431. .173609,178401. .178504,
                     179266. .179568,181358. .181447,185535. .185626,

  Query Match             38.9%;  Score 155.6;  DB 5;  Length 224187;
  Best Local Similarity   74.1%;  Pred. No. 5e-42;
  Matches  197;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||| |||||||| |||||||  | ||||| ||||| | |||||||| |||
Db      44185 ACGCCTGTAATCTCAGCACTTTGGGAGGCTGAAGTGGGTGGATCTCTTGAGGTCAGGAGT 44126

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              |||||||| |||||||||||||||||||||||| |||||||||||||||| ||||
Db      44125 TCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 44066

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||||||||||    |||||| || |||||||||||||| ||||||||||||||||||
Db      44065 CAGGCATGGTGGCATGTGCCTGTAATCTCAGCTACTCAGGAGGCGGAGATTGCAGTGAGC 44006

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||||||||        |   | |   | | ||||||  ||
Db      44005 CGAGATCGCACCACCGCACTCCAGCCTGGGTGACAGAGCAAGACTTCATCTCAAAAAATA 43946

Qy        321 CAAAAACAACAACAAAAAACAAAAAA 346
              | ||| || ||  ||||| |||| |
Db      43945 AATAAATAAAAAATAAAAATAAAATA 43920


RESULT 38
HSJ193N13
LOCUS       HSJ193N13             122961 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP1-193N13 on chromosome 6q21-22.31
            Contains the 5' end of the MAN1A1 gene for mannosidase, alpha,
            class 1A, member 1 (Man9-mannosidase (MAN9), HUMM9). Contains a CpG
            island, complete sequence.
ACCESSION   AL078600
VERSION     AL078600.15  GI:6273536
KEYWORDS    HTG; CpG island; HUMM9; MAN1A1; MAN9; Man9-mannosidase;
            mannosidase.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 122961)
  AUTHORS   Dunn,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 7, 1999 this sequence version replaced gi:6165363.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP1-193N13 is from the library RPCI-1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .122961
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="q21-22.31"
                     /clone="RP1-193N13"
                     /clone_lib="RPCI-1"
     misc_feature    1
                     /note="Clone_left_end: RP1-193N13"
     gene            join(complement(5197. .52970),
                     complement(AL022722.1:19522. .133076))
                     /gene="MAN1A1"
                     /locus_tag="RP3-354N19.2-001"
     mRNA            join(complement(52751. .52970),complement(51672. .52496),
                     complement(10105. .10201),complement(5197. .5312),
                     complement(AL022722.1:132996. .133076),
                     complement(AL022722.1:90573. .90667),
                     complement(AL022722.1:47072. .47195),
                     complement(AL022722.1:43567. .43660),
                     complement(AL022722.1:36090. .36205),
                     complement(AL022722.1:31977. .32196),
                     complement(AL022722.1:30718. .30890),
                     complement(AL022722.1:22625. .22740),
                     complement(AL022722.1:19522. .22258))
                     /gene="MAN1A1"
                     /locus_tag="RP3-354N19.2-001"
                     /note="match: ESTs: AA317167 AI122364 AI256535 AI745265
                     AL135738 AL673800 AU119410 AU135696 AV650705 AV689108
                     AV701374 BE083222 BE815845 BE877565 BG428373 BG548683
                     BI441201 BI964041 BI964247 BM538709 BM806736 BM933828
                     match: cDNAs: AF027156 AK025599 U03457 U03458 U04299
                     U04301 X74837 Y12503"
     CDS             join(complement(51672. .52274),complement(10105. .10201),
                     complement(5197. .5312),
                     complement(AL022722.1:132996. .133076),
                     complement(AL022722.1:90573. .90667),
                     complement(AL022722.1:47072. .47195),
                     complement(AL022722.1:43567. .43660),
                     complement(AL022722.1:36090. .36205),
                     complement(AL022722.1:31977. .32196),
                     complement(AL022722.1:30718. .30890),
                     complement(AL022722.1:22625. .22740),
                     complement(AL022722.1:22132. .22258))
                     /gene="MAN1A1"
                     /locus_tag="RP3-354N19.2-001"
                     /standard_name="OTTHUMP00000017107"
                     /note="match: proteins: AAH15265 CE08947 O02773 O60476
                     P33908 P39098 P45700 P45701 P53624 P53625 Q18788 Q9NR34
                     Q9VAD6"
                     /codon_start=1
                     /protein_id="CAB75695.2"
                     /db_xref="GI:56203652"
                     /db_xref="GOA:P33908"
                     /db_xref="HGNC:6821"
                     /db_xref="HSSP:1FO3"
                     /db_xref="UniProtKB/Swiss-Prot:P33908"
                     /translation="MPVGGLLPLFSSPAGGVLGGGLGGGGGRKGSGPAALRLTEKFVL
                     LLVFSAFITLCFGAIFFLPDSSKLLSGVLFHSSPALQPAADHKPGPGARAEDAAEGRA
                     RRREEGAPGDPEAALEDNLARIRENHERALREAKETLQKLPEEIQRDILLEKKKVAQD
                     QLRDKAPFRGLPPVDFVPPIGVESREPADAAIREKRAKIKEMMKHAWNNYKGYAWGLN
```

```
                     ELKPISKGGHSSSLFGNIKGATIVDALDTLFIMEMKHEFEEAKSWVEENLDFNVNAEI
                     SVFEVNIRFVGGLLSAYYLSGEEIFRKKAVELGVKLLPAFHTPSGIPWALLNMKSGIG
                     RNWPWASGGSSILAEFGTLHLEFMHLSHLSGNPIFAEKVMNIRTVLNKLEKPQGLYPN
                     YLNPSSGQWGQHHVSVGGLGDSFYEYLLKAWLMSDKTDLEAKKMYFDAVQAIETHLIR
                     KSSSGLTYIAEWKGGLLEHKMGHLTCFAGGMFALGADAAPEGMAQHYLELGAEIARTC
                     HESYNRTFMKLGPEAPRFDGGVEAIATRQNEKYYILRPEVMETYMYMWRLTHDPKYRK
                     WAWEAVEALENHCRVNGGYSGLRDVYLLHESYDDVQQSFFLAETLKYLYLIFSDDDLL
                     PLEHWIFNSEAHLLPILPKDKKEVEIREE"
     misc_feature    122961
                     /note="Clone_right_end: RP1-193N13"
ORIGIN

  Query Match          38.9%;  Score 155.4;  DB 5;  Length 122961;
  Best Local Similarity   74.7%;  Pred. No. 4.4e-42;
  Matches  210;  Conservative    1;  Mismatches   62;  Indels    8;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | ||||||||||||||| ||| |||||||||||||||| |||||||||||||||| |||
Db       35762 ACGCCTGTAATCCCAGCGCTTTGGGAGGCCAAGGTGGGAGGATCACCTGAGGTCAGGAGT 35821

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               | ||||||| |||||||||||||||||||||  |||| ||| |||||||||||||| ||||
Db       35822 TGGAGACCAGCCTGGCCAACATGGTGAAACCTGGTCTCTACCAAAAATACAAAAATTAGC 35881

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||| |||||||   |||||||| |||||||||||  ||| | |||  |||||||||||
Db       35882 TGGGCGTGGTGGCGGGCACCTGTAATCCCAGCTACTCCAGAGGCTGAGGCTGCAGTGAGC 35941

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               |||||||       | | |   |||       | | | ||  ||||  | ||:|| |||
Db       35942 TGAGATC--------ATGCCACTGCACTCCAGCCTGGGCGACAAGAGCAAAACTCCATCT 35993

Qy         321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGT 361
               ||| || || || ||||| | ||   |  |  | |  || |
Db       35994 CAAGAAAAAAAAAAAAAAGCCAATTGAGGAATATAATTTTT 36034


RESULT 39
AC004876
LOCUS       AC004876               126462 bp    DNA     linear   PRI 21-DEC-1999
DEFINITION  Homo sapiens PAC clone RP4-747G18 from 7q22-31.1, complete
            sequence.
ACCESSION   AC004876
VERSION     AC004876.2  GI:4508148
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 126462)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
   PUBMED   9847074
REFERENCE   2  (bases 1 to 126462)
  AUTHORS   Cordes,M. and Doela,D.
  TITLE     The sequence of Homo sapiens PAC clone RP4-747G18
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 126462)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 126462)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAR-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 126462)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUL-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 126462)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-DEC-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Mar 24, 1999 this sequence version replaced gi:3213121.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_DJ0747G18
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            The sequence of this clone was established as part of a mapping and
            sequencing collaboration between the NHGRI Chromosome 7 Mapping
            Project (Eric D. Green, Director), John D. McPherson in the
            Department of Genetics (Washington University), and the Washington
            University Genome Sequencing Center.  For additional information
            about the map position of this sequence, see
            http://www.nhgri.nih.gov/DIR/GTB/CHR7 , send
            mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            This clone was derived from human PAC library RPCI-4, prepared by
            Pieter de Jong and coworkers at the Roswell Park Cancer Institute
            (http://bacpac.med.buffalo.edu) using the method described by
            Ioannou et al., Nature Genetics 6:84-9 (1994).  The library is from
            one male donor.
            The clone may be obtained either from Genome Systems, Inc.
            (http://www.genomesystems.com) or Research Genetics, Inc.
            (http://www.resgen.com); or from Pieter de Jong.
            VECTOR:  pCYPAC2
            NEIGHBORING SEQUENCE INFORMATION:
            Actual start of this clone is at base position 1 of RP4-747G18;
            actual end is at 126462 of RP4-747G18.

            The double stranding from 79862 to 80357 in the PAC RP4-747G18 can
            not be guaranteed although their sequence fidelity is believed to
            be correct.  Digests using EcorV and HindIII are consistent with
            the submitted assembly.
FEATURES             Location/Qualifiers
     source          1. .126462
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q22-31.1"
                     /clone="RP4-747G18"
                     /clone_lib="RPCI-4"
     repeat_region   1. .245
                     /rpt_family="Alu"
     repeat_region   578. .898
                     /rpt_family="Alu"
     repeat_region   1238. .1373
```

```
                /rpt_family="MIR"
repeat_region   1378. .1501
                /rpt_family="Alu"
repeat_region   1502. .1796
                /rpt_family="Alu"
repeat_region   1797. .1978
                /rpt_family="Alu"
repeat_region   1981. .2265
                /rpt_family="Alu"
repeat_region   2267. .2298
                /rpt_family="(TAAA)n"
repeat_region   2828. .3139
                /rpt_family="Alu"
repeat_region   3142. .3437
                /rpt_family="Alu"
repeat_region   3544. .3602
                /rpt_family="Alu"
repeat_region   4020. .4325
                /rpt_family="Alu"
repeat_region   4409. .4437
                /rpt_family="(CA)n"
repeat_region   5668. .5956
                /rpt_family="Alu"
repeat_region   6107. .6225
                /rpt_family="MIR"
repeat_region   6226. .6575
                /rpt_family="Alu"
repeat_region   6576. .6600
                /rpt_family="MIR"
repeat_region   6601. .6912
                /rpt_family="Alu"
repeat_region   6913. .7014
                /rpt_family="MIR"
repeat_region   7066. .7340
                /rpt_family="Alu"
repeat_region   7744. .8066
                /rpt_family="Alu"
repeat_region   8662. .8875
                /rpt_family="Alu"
repeat_region   8881. .8996
                /rpt_family="Other"
repeat_region   8997. .9292
                /rpt_family="Alu"
repeat_region   9293. .9460
                /rpt_family="Other"
repeat_region   9469. .9753
                /rpt_family="Alu"
repeat_region   9754. .9783
                /rpt_family="(TAA)n"
repeat_region   9794. .10112
                /rpt_family="Alu"
repeat_region   10908. .10937
                /rpt_family="Alu"
repeat_region   10938. .10970
                /rpt_family="(TAAA)n"
repeat_region   10971. .11246
                /rpt_family="Alu"
repeat_region   11247. .11553
                /rpt_family="Alu"
repeat_region   11785. .11805
                /rpt_family="AT_rich"
repeat_region   11807. .12056
                /rpt_family="Alu"
repeat_region   12103. .12411
                /rpt_family="Alu"
misc_feature    12404. .12935
                /note="match to EST AA577797 (NID:g2355981) nn24b08.s1"
repeat_region   12412. .12635
                /rpt_family="MIR"
repeat_region   12767. .12816
                /rpt_family="Alu"
repeat_region   12820. .13088
                /rpt_family="Alu"
repeat_region   13793. .13904
                /rpt_family="Alu"
repeat_region   13905. .14214
                /rpt_family="Alu"
repeat_region   14215. .14424
                /rpt_family="Alu"
repeat_region   14720. .15036
                /rpt_family="Alu"
repeat_region   15202. .15380
                /rpt_family="Alu"
repeat_region   15383. .15684
                /rpt_family="Alu"
repeat_region   16159. .16198
                /rpt_family="MIR"
repeat_region   16219. .16337
                /rpt_family="MIR"
repeat_region   16355. .16422
                /rpt_family="polypurine"
repeat_region   16479. .16579
                /rpt_family="MIR"

  Query Match             38.9%;  Score 155.4;  DB 5;  Length 126462;
  Best Local Similarity   72.6%;  Pred. No. 4.5e-42;
  Matches  217;  Conservative    1;  Mismatches   72;  Indels    9;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||| |||||||||||||| || ||||||||||||||||||||| ||||| || ||
Db      93671 ACGCCTATAATCCCAGCACTTTGGAAGGCCAAGGTGGGCGGATCACTTGAGGCCAGGAAT 93730

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| ||||||||| |||||||||||||||| |||||||||||||||| ||||
Db      93731 TCGAGACCAGCCTGGCCAAAATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 93790

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
                ||  ||||||    | |||||||||||| |||||| |||| ||||| ||||||||||
Db      93791 CCGGAGTGGTGGTGGGCGCCTGTAGTCCCAACTACTCGGGAGGCGGAGGTTGCAGTGAGC 93850

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||| | |    |    |   | |   |   |          ||||| |||||:| ||||
Db      93851 CGAGACCACGCCTTTGCACTCCAGCCTGGGCGA---------CAGAGCGAGACTCTGTCT 93901

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGA 379
              ||| ||||| |||||||||  ||| ||    |||   ||| | |  | |    ||||
Db      93902 CAACAACAATAACAAAAAAAGAAACAAGTGCAAGGTCTTGCCTTTGAGTTGGGGTCAGA 93960


RESULT 40
AC007616/c
LOCUS       AC007616               132492 bp    DNA     linear   PRI 18-DEC-2003
DEFINITION  Homo sapiens chromosome 16 clone RP11-547D14, complete sequence.
ACCESSION   AC007616
VERSION     AC007616.5  GI:40018659
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 132492)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 132492)
  AUTHORS   Bruce,D., Mundt,M., Doggett,N., Munk,C., Saunders,E., Robinson,D.,
            Jones,M., Buckingham,J., Chasteen,L., Thompson,S., Goodwin,L.,
            Bryant,J., Tesmer,J., Meincke,L., Longmire,J., White,S., Tatum,O.,
            Campbell,C., Fawcett,J., Maltbie,M., Bussod,M., Sutherland,R.,
            McMurry,K., Han,C. and Deaven,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAY-1999) Center for Human Genome Studies, DOE Joint
            Genome Institute, Los Alamos National Laboratory, MS M888, Los
            Alamos, NM 87545, USA
```

```
REFERENCE   3  (bases 1 to 132492)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 132492)
  AUTHORS   Stanford Human Genome Center and Los Alamos National Laboratory.
  CONSRTM   DOE Joint Genome Institute
  TITLE     Direct Submission
  JOURNAL   Submitted (18-DEC-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Dec 18, 2003 this sequence version replaced gi:21700556.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center and Los Alamos
            National Laboratory
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.

            NOTE:  This is not the entire sequence of the clone (entire
            sequence 168.9kb).  It is clipped at the overlap with AC007613.
            The number of bases overlapped is 29296.
FEATURES             Location/Qualifiers
     source          1. .132492
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="RP11-547D14"
ORIGIN

  Query Match             38.9%;  Score 155.4;  DB 5;  Length 132492;
  Best Local Similarity   73.1%;  Pred. No. 4.6e-42;
  Matches  198;  Conservative    1;  Mismatches   72;  Indels    0;  Gaps    0;

Qy          83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
               ||||||||||||||||||| |||||||||||| ||| ||||||| |||||||| ||| |
Db       19740 GCCTGTAATCCCAGCACTTGGGGAGGCCAAGGCGGGTGGATCACTTGAGGTCAGGAGTTT 19681

Qy         143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
               ||||||| |||||||||| |||||||||| || | |||| |||| ||||| ||||||
Db       19680 GAGACCAGCCTGGCCAACACGGTGAAACCCCAACTCTCCTAATAATAAAAAAATTAGCTG 19621

Qy         203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
               ||| |||||||||  | |||| | |||||||||||||||||| | |||   | ||
Db       19620 GGCCTGGTGGCATGCGCCTGCAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGCT 19561

Qy         263 AGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCA 322
               || | |||   |||   ||| |       |   |     ||||| |||||:|| |||||
Db       19560 TGAACCCAGGAGGAGGGAGAATTGCTTGAACCCTGGGCGACAGAGCGAGACTCCATCTCA 19501

Qy         323 AAAACAACAACAAAAAACAAAAAAACCATAA 353
               ||||||| |  ||| || |  |||| |  ||
Db       19500 AAAACAAAACAAAACAAAACCAAAAACTAAA 19470


RESULT 41
AC027249/c
LOCUS       AC027249              145679 bp    DNA     linear   HTG 27-APR-2000
DEFINITION  Homo sapiens chromosome 11 clone RP11-753E7 map 11, WORKING DRAFT
            SEQUENCE, 35 unordered pieces.
ACCESSION   AC027249
VERSION     AC027249.2  GI:7651924
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 145679)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 11, clone RP11-753E7
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 145679)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 27, 2000 this sequence version replaced gi:7331619.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L8860
                Center clone name: 753_E_7
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 129416 bases at least Q40
                Consensus quality: 137400 bases at least Q30
                Consensus quality: 140385 bases at least Q20
                Insert size: 154000; agarose-fp
                Insert size: 142279; sum-of-contigs
                Quality coverage: 3.9 in Q20 bases; agarose-fp
                Quality coverage: 4.2 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 35 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1125: contig of 1125 bp in length
            *     1126     1225: gap of 100 bp
            *     1226     2426: contig of 1201 bp in length
            *     2427     2526: gap of 100 bp
            *     2527     3864: contig of 1338 bp in length
            *     3865     3964: gap of 100 bp
            *     3965     5223: contig of 1259 bp in length
            *     5224     5323: gap of 100 bp
            *     5324     7019: contig of 1696 bp in length
            *     7020     7119: gap of 100 bp
            *     7120     9177: contig of 2058 bp in length
            *     9178     9277: gap of 100 bp
            *     9278    11580: contig of 2303 bp in length
            *    11581    11680: gap of 100 bp
            *    11681    13580: contig of 1900 bp in length
```

```
         *    13581    13680: gap of 100 bp
         *    13681    16409: contig of 2729 bp in length
         *    16410    16509: gap of 100 bp
         *    16510    19008: contig of 2499 bp in length
         *    19009    19108: gap of 100 bp
         *    19109    21633: contig of 2525 bp in length
         *    21634    21733: gap of 100 bp
         *    21734    24035: contig of 2302 bp in length
         *    24036    24135: gap of 100 bp
         *    24136    26117: contig of 1982 bp in length
         *    26118    26217: gap of 100 bp
         *    26218    28765: contig of 2548 bp in length
         *    28766    28865: gap of 100 bp
         *    28866    31673: contig of 2808 bp in length
         *    31674    31773: gap of 100 bp
         *    31774    34537: contig of 2764 bp in length
         *    34538    34637: gap of 100 bp
         *    34638    37137: contig of 2500 bp in length
         *    37138    37237: gap of 100 bp
         *    37238    40767: contig of 3530 bp in length
         *    40768    40867: gap of 100 bp
         *    40868    44314: contig of 3447 bp in length
         *    44315    44414: gap of 100 bp
         *    44415    46751: contig of 2337 bp in length
         *    46752    46851: gap of 100 bp
         *    46852    49744: contig of 2893 bp in length
         *    49745    49844: gap of 100 bp
         *    49845    53918: contig of 4074 bp in length
         *    53919    54018: gap of 100 bp
         *    54019    57127: contig of 3109 bp in length
         *    57128    57227: gap of 100 bp
         *    57228    61720: contig of 4493 bp in length
         *    61721    61820: gap of 100 bp
         *    61821    66653: contig of 4833 bp in length
         *    66654    66753: gap of 100 bp
         *    66754    70529: contig of 3776 bp in length
         *    70530    70629: gap of 100 bp
         *    70630    75419: contig of 4790 bp in length
         *    75420    75519: gap of 100 bp
         *    75520    81973: contig of 6454 bp in length
         *    81974    82073: gap of 100 bp
         *    82074    91465: contig of 9392 bp in length
         *    91466    91565: gap of 100 bp
         *    91566    99702: contig of 8137 bp in length
         *    99703    99802: gap of 100 bp
         *    99803   107085: contig of 7283 bp in length
         *   107086   107185: gap of 100 bp
         *   107186   114948: contig of 7763 bp in length
         *   114949   115048: gap of 100 bp
         *   115049   123618: contig of 8570 bp in length
         *   123619   123718: gap of 100 bp
         *   123719   131968: contig of 8250 bp in length
         *   131969   132068: gap of 100 bp
         *   132069   145679: contig of 13611 bp in length.
FEATURES             Location/Qualifiers
   source            1. .145679
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11"
                     /clone="RP11-753E7"
                     /clone_lib="RPCI-11 Human Male BAC"
   misc_feature      1. .1125
                     /note="assembly_fragment"
   gap               1126. .1225
                     /estimated_length=100
   misc_feature      1226. .2426
                     /note="assembly_fragment"
   gap               2427. .2526
                     /estimated_length=100
   misc_feature      2527. .3864
                     /note="assembly_fragment"
   gap               3865. .3964
                     /estimated_length=100
   misc_feature      3965. .5223
                     /note="assembly_fragment"
   gap               5224. .5323
                     /estimated_length=100
   misc_feature      5324. .7019
                     /note="assembly_fragment"
   gap               7020. .7119
                     /estimated_length=100
   misc_feature      7120. .9177
                     /note="assembly_fragment"
   gap               9178. .9277
                     /estimated_length=100
   misc_feature      9278. .11580
                     /note="assembly_fragment"
   gap               11581. .11680
                     /estimated_length=100
   misc_feature      11681. .13580
                     /note="assembly_fragment"
   gap               13581. .13680
                     /estimated_length=100
   misc_feature      13681. .16409
                     /note="assembly_fragment"
   gap               16410. .16509
                     /estimated_length=100
   misc_feature      16510. .19008
                     /note="assembly_fragment"
   gap               19009. .19108
                     /estimated_length=100
   misc_feature      19109. .21633
                     /note="assembly_fragment"
   gap               21634. .21733
                     /estimated_length=100
   misc_feature      21734. .24035
                     /note="assembly_fragment"
   gap               24036. .24135
                     /estimated_length=100
   misc_feature      24136. .26117
                     /note="assembly_fragment"

  Query Match             38.9%;  Score 155.4;  DB 12;  Length 145679;
  Best Local Similarity   76.0%;  Pred. No. 4.8e-42;
  Matches  203;  Conservative    1;  Mismatches   62;  Indels    1;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||||||| ||| ||||||| |||||||| |||
Db       8524 ATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGGGGGTGGATCACTTGAGGTCAGGAGT 8465

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              |||||| ||  || |||||||||||||||||  |||| ||||||||||||||| | ||||
Db       8464 TCGAGATCAGTCTAGCCAACATGGTGAAACCTTGTCTCTACTAAAAATACAAAGATTAGC 8405

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||  ||||||| ||| |||||||||||||||||||||||| | ||||  | ||
Db       8404 TGGGTGTGGTGGCCCACGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGACAGGAGAATCG 8345

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAG-CAGAGTGAGACKCCGTC 319
              || | |||   | |  |    ||  || |  |||  | | |||||  ||| :|||||
Db       8344 CTTGAACCCAGGAGGCGGAGGCTGCAGTGAGCCAAGATCGCGACAGAGCAAGATTCCGTC 8285

Qy        320 TCAAAAACAACAACAAAAAACAAAAAA 346
              ||||||| || || ||||||   | ||
Db       8284 TCAAAAAAAAAAAAAAAAAAAGGAGAA 8258


RESULT 42
AP001024
LOCUS       AP001024              183444 bp    DNA     linear   PRI 02-JUL-2003
DEFINITION  Homo sapiens genomic DNA, chromosome 11 clone:RP11-832N8, complete
            sequence.
ACCESSION   AP001024
VERSION     AP001024.6  GI:32400666
```

```
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA
  JOURNAL   Published Only in Database (2000)
REFERENCE   2  (bases 1 to 183444)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JAN-2000) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     On Jul 1, 2003 this sequence version replaced gi:31790714.
FEATURES             Location/Qualifiers
     source          1. .183444
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q"
                     /clone="RP11-832N8"
ORIGIN

  Query Match             38.9%;  Score 155.4;  DB 5;  Length 183444;
  Best Local Similarity   76.0%;  Pred. No. 5.4e-42;
  Matches  203;  Conservative    1;  Mismatches   62;  Indels    1;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               ||||||||||||||||||||| |||||||||||| ||| ||||||| |||||||| |||
Db      173708 ATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGGGGGTGGATCACTTGAGGTCAGGAGT 173767

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               |||||| ||  || |||||||||||||||||  |||| ||||||||||||||| | ||||
Db      173768 TCGAGATCAGTCTAGCCAACATGGTGAAACCTTGTCTCTACTAAAAATACAAAGATTAGC 173827

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||  ||||||| ||| |||||||||||||||||||||||| | ||||  | ||
Db      173828 TGGGTGTGGTGGCCCACGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGACAGGAGAATCG 173887

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAG-CAGAGTGAGACKCCGTC 319
               || | |||   | |  |    ||  || |  |||  | | |||||  ||| :|||||
Db      173888 CTTGAACCCAGGAGGCGGAGGCTGCAGTGAGCCAAGATCGCGACAGAGCAAGATTCCGTC 173947

Qy         320 TCAAAAACAACAACAAAAAACAAAAAA 346
               ||||||| || || ||||||   | ||
Db      173948 TCAAAAAAAAAAAAAAAAAAAGGAGAA 173974


RESULT 43
AC018423/c
LOCUS       AC018423              202138 bp    DNA     linear   HTG 06-MAY-2001
DEFINITION  Homo sapiens chromosome 11 clone RP11-531E6 map 11, WORKING DRAFT
            SEQUENCE, 22 unordered pieces.
ACCESSION   AC018423
VERSION     AC018423.4  GI:10045404
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 202138)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 11, clone RP11-531E6
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 202138)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Castle,A., Colangelo,M.,
            Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K.,
            Domino,M., Doyle,M., Fenestor,J., Ferreira,P., FitzHugh,W.,
            Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B.,
            Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L.,
            Karatas,A., Klein,J., Landers,T., Lehoczky,J., Lieu,C., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T.,
            O'Donnell,P., Peterson,K., Pierre,N., Pollara,V., Riley,R.,
            Rothman,D., Roy,A., Santos,R., Severy,P., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 9, 2000 this sequence version replaced gi:6649385.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L3744
                Center clone name: 531_E_6
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 188769 bases at least Q40
                Consensus quality: 195790 bases at least Q30
                Consensus quality: 198059 bases at least Q20
                Insert size: 210000; agarose-fp
                Insert size: 200038; sum-of-contigs
                Quality coverage: 4.4 in Q20 bases; agarose-fp
                Quality coverage: 4.6 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 22 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2851: contig of 2851 bp in length
            *     2852     2951: gap of 100 bp
            *     2952     4977: contig of 2026 bp in length
            *     4978     5077: gap of 100 bp
            *     5078     8300: contig of 3223 bp in length
            *     8301     8400: gap of 100 bp
            *     8401    10751: contig of 2351 bp in length
            *    10752    10851: gap of 100 bp
            *    10852    14527: contig of 3676 bp in length
            *    14528    14627: gap of 100 bp
            *    14628    53554: contig of 38927 bp in length
            *    53555    53654: gap of 100 bp
            *    53655    59471: contig of 5817 bp in length
            *    59472    59571: gap of 100 bp
            *    59572    65624: contig of 6053 bp in length
            *    65625    65724: gap of 100 bp
            *    65725    72213: contig of 6489 bp in length
            *    72214    72313: gap of 100 bp
            *    72314    76349: contig of 4036 bp in length
            *    76350    76449: gap of 100 bp
            *    76450    82044: contig of 5595 bp in length
            *    82045    82144: gap of 100 bp
```

```
            *    82145    90316: contig of 8172 bp in length
            *    90317    90416: gap of 100 bp
            *    90417    96732: contig of 6316 bp in length
            *    96733    96832: gap of 100 bp
            *    96833   106557: contig of 9725 bp in length
            *   106558   106657: gap of 100 bp
            *   106658   114415: contig of 7758 bp in length
            *   114416   114515: gap of 100 bp
            *   114516   122744: contig of 8229 bp in length
            *   122745   122844: gap of 100 bp
            *   122845   132174: contig of 9330 bp in length
            *   132175   132274: gap of 100 bp
            *   132275   144490: contig of 12216 bp in length
            *   144491   144590: gap of 100 bp
            *   144591   155574: contig of 10984 bp in length
            *   155575   155674: gap of 100 bp
            *   155675   173876: contig of 18202 bp in length
            *   173877   173976: gap of 100 bp
            *   173977   199611: contig of 25635 bp in length
            *   199612   199711: gap of 100 bp
            *   199712   202138: contig of 2427 bp in length.
FEATURES             Location/Qualifiers
     source          1. .202138
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11"
                     /clone="RP11-531E6"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .2851
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     gap             2852. .2951
                     /estimated_length=100
     misc_feature    2952. .4977
                     /note="assembly_fragment"
     gap             4978. .5077
                     /estimated_length=100
     misc_feature    5078. .8300
                     /note="assembly_fragment"
     gap             8301. .8400
                     /estimated_length=100
     misc_feature    8401. .10751
                     /note="assembly_fragment"
     gap             10752. .10851
                     /estimated_length=100
     misc_feature    10852. .14527
                     /note="assembly_fragment"
     gap             14528. .14627
                     /estimated_length=100
     misc_feature    14628. .53554
                     /note="assembly_fragment"
     gap             53555. .53654
                     /estimated_length=100
     misc_feature    53655. .59471
                     /note="assembly_fragment"
     gap             59472. .59571
                     /estimated_length=100
     misc_feature    59572. .65624
                     /note="assembly_fragment"
     gap             65625. .65724
                     /estimated_length=100
     misc_feature    65725. .72213
                     /note="assembly_fragment"
     gap             72214. .72313
                     /estimated_length=100
     misc_feature    72314. .76349
                     /note="assembly_fragment"
     gap             76350. .76449
                     /estimated_length=100
     misc_feature    76450. .82044
                     /note="assembly_fragment"
     gap             82045. .82144
                     /estimated_length=100
     misc_feature    82145. .90316
                     /note="assembly_fragment"
     gap             90317. .90416
                     /estimated_length=100
     misc_feature    90417. .96732
                     /note="assembly_fragment"
     gap             96733. .96832
                     /estimated_length=100
     misc_feature    96833. .106557
                     /note="assembly_fragment"
     gap             106558. .106657
                     /estimated_length=100
     misc_feature    106658. .114415
                     /note="assembly_fragment"
     gap             114416. .114515
                     /estimated_length=100
     misc_feature    114516. .122744
                     /note="assembly_fragment"
     gap             122745. .122844
                     /estimated_length=100
     misc_feature    122845. .132174
                     /note="assembly_fragment"
     gap             132175. .132274
                     /estimated_length=100
     misc_feature    132275. .144490
                     /note="assembly_fragment"
     gap             144491. .144590
                     /estimated_length=100
     misc_feature    144591. .155574
                     /note="assembly_fragment"
     gap             155575. .155674
                     /estimated_length=100
     misc_feature    155675. .173876
                     /note="assembly_fragment"
     gap             173877. .173976
                     /estimated_length=100

  Query Match           38.9%;  Score 155.4;  DB 12;  Length 202138;
  Best Local Similarity   76.0%;  Pred. No. 5.6e-42;
  Matches  203;  Conservative    1;  Mismatches   62;  Indels    1;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||||||| ||| ||||||| |||||||| ||
Db     102362 ATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGGGGGTGGATCACTTGAGGTCAGGAGT 102303

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              |||||| ||  || |||||||||||||||||  |||| |||||||||||||| | ||||
Db     102302 TCGAGATCAGTCTAGCCAACATGGTGAAACCTTGTCTCTACTAAAAATACAAAGATTAGC 102243

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||  ||||||| ||| |||||||||||||||||||||||| | ||||  | ||
Db     102242 TGGGTGTGGTGGCCCACGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGACAGGAGAATCG 102183

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAG-CAGAGTGAGACKCCGTC 319
                || | |||   | |  |    ||  || |  |||  | | |||||  ||| :|||||
Db     102182 CTTGAACCCAGGAGGCGGAGGCTGCAGTGAGCCAAGATCGCGACAGAGCAAGATTCCGTC 102123

Qy        320 TCAAAAACAACAACAAAAAACAAAAAA 346
              ||||||| || || ||||||   | ||
Db     102122 TCAAAAAAAAAAAAAAAAAAAGGAGAA 102096


RESULT 44
AC099489
LOCUS       AC099489              204493 bp    DNA     linear   PRI 04-FEB-2003
DEFINITION  Homo sapiens chromosome 16 clone CTD-3088G3, complete sequence.
ACCESSION   AC099489
VERSION     AC099489.2  GI:28201477
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 204493)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 204493)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 204493)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-FEB-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Feb 4, 2003 this sequence version replaced gi:16930905.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center and Los Alamos
            National Laboratory
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.8% of Sequence;
            Estimated Total Number of Errors is 0.4.
            NOTE:  Shatter libraries failed to verify the dinucleotide repeat
            region   65514-65789.  Unsure number of repeat copies 65514-65789.
            Forced join at 65675.
FEATURES             Location/Qualifiers
     source          1. .204493
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="CTD-3088G3"
     misc_feature    65514. .65789
                     /note="NOTE:  Shatter libraries failed to verify the
                     dinucleotide repeat region  65514-65789.  Unsure number of
                     repeat copies 65514-65789.  Forced join at 65675."
ORIGIN

  Query Match             38.9%;  Score 155.4;  DB 5;  Length 204493;
  Best Local Similarity   73.1%;  Pred. No. 5.7e-42;
  Matches  198;  Conservative    1;  Mismatches   72;  Indels    0;  Gaps    0;

Qy          83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
               ||||||||||||||||||| |||||||||||| ||| ||||||| |||||||| ||| |
Db       14479 GCCTGTAATCCCAGCACTTGGGGAGGCCAAGGCGGGTGGATCACTTGAGGTCAGGAGTTT 14538

Qy         143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
               ||||||| |||||||||| |||||||||||  || | |||| |||| ||||| ||||||
Db       14539 GAGACCAGCCTGGCCAACACGGTGAAACCCCAACTCTCCTAATAATAAAAAAATTAGCTG 14598

Qy         203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
               ||| |||||||||  | |||| | |||||||||||||||||| | |||   | ||
Db       14599 GGCCTGGTGGCATGCGCCTGCAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGCT 14658

Qy         263 AGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCA 322
               || | |||   |||    ||| |       |   |     ||||| |||||:|| |||||
Db       14659 TGAACCCAGGAGGAGGGAGAATTGCTTGAACCCTGGGCGACAGAGCGAGACTCCATCTCA 14718

Qy         323 AAAACAACAACAAAAAACAAAAAAACCATAA 353
               ||||||| |  ||| || |  |||| |  ||
Db       14719 AAAACAAAACAAAACAAAACCAAAAACTAAA 14749


RESULT 45
AL355388
LOCUS       AL355388               205463 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-336K24 on chromosome 1 Contains
            the 5' end of the RIT1 gene for Ras-like without CAAX 1, the gene
            for a novel protein (KIAA0907), the ARHGEF2 gene for rho/rac
            guanine nucleotide exchange factor (GEF) 2, four novel genes, the
            SSR2 gene for signal sequence receptor beta (translocon-associated
            protein beta), the C1orf6 gene for chromosome 1 open reading frame
            6, the gene for mitogen-activated protein-binding
            protein-interacting protein (MAPBPIP), the RAB25 gene for RAB25
            (member RAS oncogene family), the 5' end of the LMNA gene for lamin
            A/C and three CpG islands, complete sequence.
ACCESSION   AL355388
VERSION     AL355388.30  GI:29367464
KEYWORDS    HTG; ARHGEF2; C1orf6; KIAA0907; LMNA; MAPBPIP; RAB25; RIT1; SSR2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 205463)
  AUTHORS   Hall,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 29, 2003 this sequence version replaced gi:28446036.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-336K24 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            ---------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            ----------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .205463
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-336K24"
                     /clone_lib="RPCI-11.2"
     misc_feature    1
                     /note="Clone_left_end: RP11-336K24"
     gene            complement(join(3775. .4186,4336. .4471))
                     /gene="RIT1"
                     /locus_tag="RP11-101O6.4-004"
     mRNA            complement(join(3775. .4186,4336. .4471))
                     /gene="RIT1"
                     /locus_tag="RP11-101O6.4-004"
                     /product="Ras-like without CAAX 1"
                     /note="match: ESTs: Em:BX644870.1"
     gene            join(complement(4923. .5084),complement(4336. .4484),
                     complement(4130. .4186),
                     complement(AL139128.24:141244. .141317),
                     complement(AL139128.24:140824. .141015),
                     complement(AL139128.24:134321. .137131))
```

```
                     /gene="RIT1"
                     /locus_tag="RP11-101O6.4-001"
     mRNA            join(complement(4923. .5084),complement(4336. .4484),
                     complement(4130. .4186),
                     complement(AL139128.24:141244. .141317),
                     complement(AL139128.24:140824. .141015),
                     complement(AL139128.24:134321. .137131))
                     /gene="RIT1"
                     /locus_tag="RP11-101O6.4-001"
                     /product="Ras-like without CAAX 1"
                     /note="match: ESTs: Em:BF185015.1 Em:BF208489.1
                     match: cDNAs: Em:AF084462.1 Em:U71203.1 Em:U78165.1
                     Em:Y07566.1"
     gene            join(complement(4923. .5024),complement(4336. .4484),
                     complement(4130. .4186),
                     complement(AL139128.24:141244. .141317),
                     complement(AL139128.24:141116. .141123),
                     complement(AL139128.24:140824. .141015),
                     complement(AL139128.24:136911. .137131))
                     /gene="RIT1"
                     /locus_tag="RP11-101O6.4-003"
     mRNA            join(complement(4923. .5024),complement(4336. .4484),
                     complement(4130. .4186),
                     complement(AL139128.24:141244. .141317),
                     complement(AL139128.24:141116. .141123),
                     complement(AL139128.24:140824. .141015),
                     complement(AL139128.24:136911. .137131))
                     /gene="RIT1"
                     /locus_tag="RP11-101O6.4-003"
                     /product="Ras-like without CAAX 1"
                     /note="match: ESTs: Em:BE782613.1"
     gene            join(complement(4558. .4595),complement(4336. .4484),
                     complement(4130. .4186),
                     complement(AL139128.24:141244. .141317),
                     complement(AL139128.24:140824. .141015),
                     complement(AL139128.24:136787. .137131))
                     /gene="RIT1"
                     /locus_tag="RP11-101O6.4-002"
     mRNA            join(complement(4558. .4595),complement(4336. .4484),
                     complement(4130. .4186),
                     complement(AL139128.24:141244. .141317),
                     complement(AL139128.24:140824. .141015),
                     complement(AL139128.24:136787. .137131))
                     /gene="RIT1"
                     /locus_tag="RP11-101O6.4-002"
                     /product="Ras-like without CAAX 1"
                     /note="match: ESTs: Em:BQ918446.1"
     CDS             join(complement(4558. .4565),complement(4336. .4484),
                     complement(4130. .4186),
                     complement(AL139128.24:141244. .141317),
                     complement(AL139128.24:140824. .141015),
                     complement(AL139128.24:136901. .137131))
                     /gene="RIT1"
                     /locus_tag="RP11-101O6.4-002"
                     /standard_name="OTTHUMP00000015933"
                     /codon_start=1
                     /product="Ras-like without CAAX 1"
                     /protein_id="CAH72622.1"
                     /db_xref="GI:55663544"
                     /db_xref="GOA:Q5VY89"
                     /db_xref="InterPro:IPR001806"
                     /db_xref="InterPro:IPR002041"
                     /db_xref="InterPro:IPR003577"
                     /db_xref="InterPro:IPR003578"
                     /db_xref="InterPro:IPR003579"
                     /db_xref="InterPro:IPR005225"
                     /db_xref="UniProtKB/TrEMBL:Q5VY89"
                     /translation="MERWLFLGATEEGPKRTMDSGTRPVGSCCSSPAGLSREYKLVML
                     GAGGVGKSAMTMQFISHRFPEDHDPTIEDAYKIRIRIDDEPANLDILDTAGQAEFTAM
                     RDQYMRAGEGFIICYSITDRRSFHEVREFKQLIYRVRRTDDTPVVLVGNKSDLKQLRQ
                     VTKEEGLALAREFSCPFFETSAAYRYYIDDVFHALVREIRRKEKEAVLAMEKKSKPKN
                     SVWKRLKSPFRKKKDSVT"
     CDS             join(complement(4336. .4441),complement(4130. .4186),
                     complement(AL139128.24:141244. .141317),
                     complement(AL139128.24:140824. .141015),
                     complement(AL139128.24:136901. .137131))
                     /gene="RIT1"
                     /locus_tag="RP11-101O6.4-001"
                     /standard_name="OTTHUMP00000015932"
                     /note="match: proteins: Sw:P70426 Sw:Q92963"
                     /codon_start=1
                     /product="Ras-like without CAAX 1"
                     /protein_id="CAH72621.1"
                     /db_xref="GI:55663543"
                     /db_xref="InterPro:IPR001806"
                     /db_xref="InterPro:IPR005225"
                     /translation="MDSGTRPVGSCCSSPAGLSREYKLVMLGAGGVGKSAMTMQFISH
                     RFPEDHDPTIEDAYKIRIRIDDEPANLDILDTAGQAEFTAMRDQYMRAGEGFIICYSI
                     TDRRSFHEVREFKQLIYRVRRTDDTPVVLVGNKSDLKQLRQVTKEEGLALAREFSCPF
                     FETSAAYRYYIDDVFHALVREIRRKEKEAVLAMEKKSKPKNSVWKRLKSPFRKKKDSV
                     T"
     gene            complement(join(6723. .8000,9553. .9644,10305. .10417,
                     11179. .11352,15055. .15308,15541. .15598,17307. .17367,
                     19312. .19523,20356. .20519,20784. .20836,22976. .23055,
                     23392. .23520,27313. .27529,28016. .28077))
                     /gene="RP11-336K24.1"
                     /locus_tag="RP11-336K24.1-001"
     mRNA            complement(join(6723. .8000,9553. .9644,10305. .10417,
                     11179. .11352,15055. .15308,15541. .15598,17307. .17367,
                     19312. .19523,20356. .20519,20784. .20836,22976. .23055,
                     23392. .23520,27313. .27529,28016. .28077))
                     /gene="RP11-336K24.1"
                     /locus_tag="RP11-336K24.1-001"
                     /product="novel protein"
                     /note="match: cDNAs: Em:AK012969.1 Em:AY112680.1
                     Em:BC006621.1 Em:BC062637.1"
     polyA_site      complement(6723)
                     /gene="RP11-336K24.1"
                     /locus_tag="RP11-336K24.1-001"
     gene            complement(join(6725. .9644,10305. .10417,11179. .11352,
                     15055. .15308,15541. .15598,17307. .17367,19312. .19523,
                     20356. .20519,20784. .20836,22976. .23055,23392. .23520,
                     27313. .27529,28016. .28079))
                     /gene="RP11-336K24.1"
                     /locus_tag="RP11-336K24.1-003"
     mRNA            complement(join(6725. .9644,10305. .10417,11179. .11352,
                     15055. .15308,15541. .15598,17307. .17367,19312. .19523,
                     20356. .20519,20784. .20836,22976. .23055,23392. .23520,
                     27313. .27529,28016. .28079))
                     /gene="RP11-336K24.1"
                     /locus_tag="RP11-336K24.1-003"
                     /product="novel protein"
                     /note="match: cDNAs: Em:AB020714.1 Em:AY112681.1"
     polyA_site      complement(6725)
                     /gene="RP11-336K24.1"

  Query Match           38.9%;  Score 155.4;  DB 5;  Length 205463;
  Best Local Similarity  74.7%;  Pred. No. 5.7e-42;
  Matches  207;  Conservative    1;  Mismatches   67;  Indels    2;  Gaps    1;

Qy        81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
             |||||||||||||||||||||||||||||  ||| ||| ||||||||||||||||| |||
Db    177219 ACGCCTGTAATCCCAGCACTTCGGGAGGCAGAGGCGGGTGGATCACCTGAGGTCAGGAGT 177278

Qy       141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
             |||||||| |||| |||||| |||||||| ||||| ||||||||||||||||| ||||
Db    177279 TCGAGACCAGCCTGACCAACAAGGTGAAACCTCGTCTCTACTAAAAATACAAAAATTAGC 177338

Qy       201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
             | |||||||||| || |||||||||||||||||||||||| | ||||  |  | ||
Db    177339 CGAGCATGGTGGCAGACGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGACAG--GAGAAT 177396

Qy       261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
             ||      | · |   || ||| |  |||       ||   ||||||||||||:|||||
Db    177397 TGCTTGAACCCGGGAGGCGGAAGTTGCAGTGAGTCTAGATCGCAGAGTGAGACTCCGTCC 177456
```

```
Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
              |    |||| || | ||||  |||| |    || | |
Db     177457 CCCCCACAAAAAAACAAAAACAAAACAAACAAACAAA 177493


RESULT 46
AC026423
LOCUS       AC026423              116130 bp    DNA     linear   PRI 18-MAR-2003
DEFINITION  Homo sapiens chromosome 16 clone CTD-2169M19, complete sequence.
ACCESSION   AC026423
VERSION     AC026423.9  GI:29029225
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 116130)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 116130)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 116130)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JAN-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 116130)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Mar 18, 2003 this sequence version replaced gi:18057078.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center and Los Alamos
            National Laboratory
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
FEATURES             Location/Qualifiers
     source          1. .116130
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="CTD-2169M19"
ORIGIN

  Query Match             38.8%;  Score 155.2;  DB 5;  Length 116130;
  Best Local Similarity   74.5%;  Pred. No. 5.1e-42;
  Matches  210;  Conservative    1;  Mismatches   64;  Indels    7;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | | |||||||||||||||||| |||||||||||||||| ||||||| ||||| || |||
Db      11307 ACGTCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGTGGATCACTTGAGGCCAGGAGT 11366

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||||||||||||||||| ||| | ||| ||||||||||||||||| ||||
Db      11367 TCGAGACCAGCCTGGCCAACATGGTGAGACCTCATCTCTACTAAAAATACAAAAATTAGC 11426

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||| ||||||||    |||||| |||||||||||||||| | |||   | ||  |
Db      11427 CGGGCTTGGTGGCATGTGCCTGTAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAAGGT 11486

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGAT-------CACAGAGTGAGCAGAGTGAGAC 313
              ||| | | || | |    ||  |||   |         |   | ||||||||||||||


Db      11487 GGAGGTTGTAGTGAGCCGAGATTGCACCACTGCACTCCAGCCTGGGCAGCAGAGTGAGAC 11546

Qy        314 KCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGA 355
              : |||||||||| || || |||||| |||||||  | || |
Db      11547 TTGGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAAAA 11588


RESULT 47
AC079456/c
LOCUS       AC079456              177640 bp    DNA     linear   PRI 22-OCT-2002
DEFINITION  Homo sapiens 12 BAC RP11-206B11 (Roswell Park Cancer Institute
            Human BAC Library) complete sequence.
ACCESSION   AC079456
VERSION     AC079456.28  GI:24211071
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 177640)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S.,
            Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P.,
            Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N.,
            Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P.,
            Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M.,
            Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S.,
            Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y.,
            Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U.,
            King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N.,
            Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O.,
            Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X.,
            Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P.,
            Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E.,
            Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S.,
            Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T.,
            Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M.,
            Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 177640)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-SEP-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 177640)
  AUTHORS   Worley,K.C.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (30-SEP-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 177640)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-OCT-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   5  (bases 1 to 177640)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   6  (bases 1 to 177640)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   7  (bases 1 to 177640)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   8  (bases 1 to 177640)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Oct 22, 2002 this sequence version replaced gi:15809129.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
               STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
               Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
               Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.

            SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
            standard of double strand coverage with a minimum of 2 clones and 2
            reads with no ambiguities or 2 chemistries with a minimum of 2
            clones and 3 reads with no ambiguities. If the sequence quality for
            a region does not meet this standard, it will be indicated in the
            annotation as Low Coverage.

            QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
            standards - estimated error rate less than 1 per 10,000 bases.
            Reports of lowest quality individual bases and measures of base
            quality are listed below. Description of the metrics can be found
            at URL:
            http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht
            ml.
FEATURES             Location/Qualifiers
     source          1. .177640
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /clone="RP11-206B11"
     repeat_region   complement(142. .457)
                     /rpt_family="L1M4"
     repeat_region   458. .665
                     /rpt_family="AluSq"
     repeat_region   complement(666. .735)
                     /rpt_family="L1M4"
     repeat_region   1056. .1084
                     /rpt_family="AT_rich"
     repeat_region   complement(1085. .2001)
                     /rpt_family="L1MB2"
     repeat_region   complement(2024. .2219)
                     /rpt_family="L1MB2"
     repeat_region   2229. .2656
                     /rpt_family="MLT1C"
     repeat_region   3478. .3792
                     /rpt_family="L2"
     repeat_region   3912. .3944
                     /rpt_family="(A)n"
     repeat_region   3965. .4097
                     /rpt_family="L1"
     repeat_region   4172. .4633
                     /rpt_family="L2"
     repeat_region   complement(5123. .5281)
                     /rpt_family="L1M4"
     repeat_region   6118. .6144
                     /rpt_family="(TTA)n"
     repeat_region   6865. .7174
                     /rpt_family="AluY"
     repeat_region   7357. .7418
                     /rpt_family="(TTTTG)n"
     repeat_region   complement(7419. .7701)
                     /rpt_family="AluSx"
     repeat_region   complement(7799. .8336)
                     /rpt_family="L1MC2"
     repeat_region   complement(8569. .8859)
                     /rpt_family="MER102"
     repeat_region   9235. .9548
                     /rpt_family="AluSx"
     repeat_region   10626. .10660
                     /rpt_family="AT_rich"
     repeat_region   complement(11325. .11685)
                     /rpt_family="L1PA13"
     repeat_region   11688. .11817
                     /rpt_family="L1PA13"
     repeat_region   12300. .12336
                     /rpt_family="(TTCA)n"
     repeat_region   complement(12572. .13010)
                     /rpt_family="LTR33"
     repeat_region   13837. .14147
                     /rpt_family="AluSx"
     repeat_region   14431. .14811
                     /rpt_family="MLT1J2"
     repeat_region   15276. .15956
                     /rpt_family="L1PA4"
     repeat_region   15959. .16015
                     /rpt_family="A-rich"
     repeat_region   16953. .17120

  Query Match             38.8%;  Score 155.2;  DB 5;  Length 177640;
  Best Local Similarity   73.7%;  Pred. No. 6.2e-42;
  Matches  196;  Conservative    1;  Mismatches   69;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| ||||||||||   || ||||||||||||||| |||
Db      97632 ATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGAAAGGTGGATCACCTGAGGTCAGGAGT 97573

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
```

```
                || ||||||  ||||||||||||||||||||||  |||| ||||||||  ||||||||  ||||
Db      97572 TCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTCTACTAAAACTACAAAAATTAGC 97513

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||  |||||  |||||||||||||||||||||||||||||| | |||   | ||   |
Db      97512 CGGCTCTGGTGACTCACACCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATGG 97453

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                || | |||   | |  |    || ||   |   |     ||||  | ||: | |||
Db      97452 CTTGAACCCAGGAGGCGGAGGTGACAGAGCGAGCCTGGGTGATAGAGCAATACTTCATCT 97393

Qy        321 CAAAAACAACAACAAAAAACAAAAAA 346
              |||||||||  |  ||| || | ||||
Db      97392 CAAAAACAAAACAAAACAAAACAAAA 97367


RESULT 48
AC068929/c
LOCUS       AC068929               178079 bp    DNA     linear   HTG 07-DEC-2000
DEFINITION  Homo sapiens chromosome 1 clone RP11-206C14 map 1, WORKING DRAFT
            SEQUENCE, 11 unordered pieces.
ACCESSION   AC068929
VERSION     AC068929.3  GI:11597096
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 178079)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 1, clone RP11-206C14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 178079)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Dec 7, 2000 this sequence version replaced gi:8671969.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L7679
                Center clone name: 206_C_14
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 42% of reads
                Sequencing vector: Plasmid; n/a; 58% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 172732 bases at least Q40
                Consensus quality: 175502 bases at least Q30
                Consensus quality: 176567 bases at least Q20
                Insert size: 170000; agarose-fp
                Insert size: 177079; sum-of-contigs
                Quality coverage: 6.2 in Q20 bases; agarose-fp
                Quality coverage: 5.9 in Q20 b.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 11 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     8548: contig of 8548 bp in length
            *     8549     8648: gap of 100 bp
            *     8649     9685: contig of 1037 bp in length
            *     9686     9785: gap of 100 bp
            *     9786    11207: contig of 1422 bp in length
            *    11208    11307: gap of 100 bp
            *    11308    14537: contig of 3230 bp in length
            *    14538    14637: gap of 100 bp
            *    14638    22686: contig of 8049 bp in length
            *    22687    22786: gap of 100 bp
            *    22787    34297: contig of 11511 bp in length
            *    34298    34397: gap of 100 bp
            *    34398    47850: contig of 13453 bp in length
            *    47851    47950: gap of 100 bp
            *    47951    67280: contig of 19330 bp in length
            *    67281    67380: gap of 100 bp
            *    67381    88352: contig of 20972 bp in length
            *    88353    88452: gap of 100 bp
            *    88453   151825: contig of 63373 bp in length
            *   151826   151925: gap of 100 bp
            *   151926   178079: contig of 26154 bp in length.
FEATURES             Location/Qualifiers
     source          1. .178079
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1"
                     /clone="RP11-206C14"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .8548
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     gap             8549. .8648
                     /estimated_length=100
     misc_feature    8649. .9685
                     /note="assembly_fragment"
     gap             9686. .9785
                     /estimated_length=100
     misc_feature    9786. .11207
                     /note="assembly_fragment"
     gap             11208. .11307
                     /estimated_length=100
     misc_feature    11308. .14537
                     /note="assembly_fragment"
     gap             14538. .14637
                     /estimated_length=100
     misc_feature    14638. .22686
                     /note="assembly_fragment"
     gap             22687. .22786
                     /estimated_length=100
     misc_feature    22787. .34297
                     /note="assembly_fragment"
     gap             34298. .34397
                     /estimated_length=100
```

```
     misc_feature    34398. .47850
                     /note="assembly_fragment"
     gap             47851. .47950
                     /estimated_length=100
     misc_feature    47951. .67280
                     /note="assembly_fragment"
     gap             67281. .67380
                     /estimated_length=100
     misc_feature    67381. .88352
                     /note="assembly_fragment"
     gap             88353. .88452
                     /estimated_length=100
     misc_feature    88453. .151825
                     /note="assembly_fragment"
     gap             151826. .151925
                     /estimated_length=100
     misc_feature    151926. .178079
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             38.8%;  Score 155.2;  DB 12;  Length 178079;
  Best Local Similarity   73.7%;  Pred. No. 6.2e-42;
  Matches  196;  Conservative    1;  Mismatches   69;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| ||||||||||||   || ||||||||||||||| |||
Db     130549 ATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGAAAGGTGGATCACCTGAGGTCAGGAGT 130490

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || |||||| |||||||||||||||||||||  |||| |||||||| |||||||| ||||
Db     130489 TCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTCTACTAAAACTACAAAAATTAGC 130430

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
                ||   ||||| | |||||||||||||||||||||||||||||| | |||   | ||   |
Db     130429 CGGCTCTGGTGACTCACACCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATGG 130370

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                || | |||   | |  |    || ||    |   |     |||| | ||: | |||
Db     130369 CTTGAACCCAGGAGGCGGAGGTGACAGAGCGAGCCTGGGTGATAGAGCAATACTTCATCT 130310

Qy        321 CAAAAACAACAACAAAAAACAAAAAA 346
              ||||||||| |  ||| || | ||||
Db     130309 CAAAAACAAAACAAAACAAAACAAAA 130284


RESULT 49
AL162384
LOCUS       AL162384               68465 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-218I7 on chromosome 9 Contains a
            novel gene similar to RAN binding protein 6 (RANBP6), the 5' end of
            a novel gene, a novel gene and a CpG island, complete sequence.
ACCESSION   AL162384
VERSION     AL162384.14  GI:13274331
KEYWORDS    HTG; CpG island; RANBP6.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 68465)
  AUTHORS   Tracey,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 12, 2001 this sequence version replaced gi:12831806.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr9
            RP11-218I7 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            ------------ Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .68465
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="9"
                     /clone="RP11-218I7"
                     /clone_lib="RPCI-11.1"
     misc_feature    1
                     /note="Clone_left_end: RP11-218I7"
     gene            join(complement(3518. .4059),
                     complement(AL353606.12:14362. .14562),
                     complement(AL365360.9:149198. .150476))
                     /locus_tag="RP11-546N22.5-001"
     mRNA            join(complement(3518. .4059),
                     complement(AL353606.12:14362. .14562),
                     complement(AL365360.9:149198. .150476))
                     /locus_tag="RP11-546N22.5-001"
                     /note="match: ESTs: AI004775.1 AI760849.1 AW370897.1
                     AW449596.1 AW629298.1 AW889233.1 BE702595.1 BE702599.1
                     BF514001.1 BF514858.1 BI079752.1 BQ376550.1
                     match: cDNAs: BC023491.1"
     gene            complement(7366. .11941)
                     /locus_tag="RP11-218I7.1-001"
     mRNA            complement(7366. .11941)
                     /locus_tag="RP11-218I7.1-001"
                     /note="match: ESTs: AI568841.1 AL513854.1 AL708378.1
                     AU132880.1 AU137263.1 AW975761.1 BF240511.1 BG109871.1
                     BG622522.1 BI091767.1 BI827617.1 BM468937.1 BM556244.1
                     BM677000.1 BM799977.1 BM800655.1 BQ001465.1 BQ068431.1
                     BQ217175.1 BQ220742.1 BQ653718.1 BQ722086.1
                     match: cDNAs: AF039023.1 AF294327.1 AK001982.1 AK017701.1
                     BC001497.1 BC012805.1 BC019309.1 BC019838.1 BC020012.1
                     BC026487.1 U72761.1 Y08890.1"
     polyA_site      complement(7366)
                     /locus_tag="RP11-218I7.1-001"
     polyA_signal    complement(7378. .7383)
                     /locus_tag="RP11-218I7.1-001"
     CDS             complement(8613. .11930)
                     /locus_tag="RP11-218I7.1-001"
                     /standard_name="OTTHUMP00000021039"
                     /note="match: proteins: AAH01497 AAH19309 O00410 O60518
                     Q96E78 Q9EQ30"
                     /codon_start=1
                     /protein_id="CAI12713.1"
                     /db_xref="GI:55957960"
                     /translation="MAATASAGVPATVSEKQEFYQLLKNLINPSCMVRRQAEEIYENI
                     PGLCKTTFLLDAVRNRRAGYEVRQMAAALLRRLLSSGFEEVYPNLPADVQRDVKIELI
                     LAVKLETHASMRKKLCDIFAVLARNLIDEDGTNHWPEGLKFLIDSIYSKNVVLWEVAL
                     HVFWHFPGIFGTQERHDLDIIKRLLDQCIQDQEHPAIRTLSARAAAAFVLANENNIAL
                     FKDFADLLPGILQAVNDSCYQDDDSVLESLVEIADTVPKYLGPYLEDTLQLSLKLCGD
                     SRLSNLQRQLALEVIVTLSETATPMLKKHTNIIAQAVPHILAMMVDLQDDEDWVNADE
```

```
                     MEEDDFDSNAVAAESALDRLACGLGGKVVLPMTKEHIMQMLQSPDWKYRHAGLMALSA
                     IGEGCHQQMESILDETVNSVLLFLQDPHPRVRAAACTTLGQMATDFAPNFQKKFHETV
                     IAALLRTMENQGNQRVQSHAASALIIFIEDCPKSLLVLYVDSMVKNLHSVLVIKLQEL
                     IRNGTKLALEQLVTTIASVADTIEEKFVPYYDIFMPSLKHIVELAVQKELKLLRGKTI
                     ECISHIGLAVGKEKFMQDASNVMQLLLKTQSDLNNMEDDDPQTSYMVSAWARMCKILG
                     KDFQQYLPLVIEPLIKTASAKPDVALLDTQDVENMSDDDGWQFVNLGDQQSFGIKTSG
                     LEAKATACQMLVYYAKELREGFVEYTEQVVKLMVPLLKFYFHDNVRVAAAESMPFLLE
                     CARIRGPEYLAQMWQFICDPLIKAIGTEPDTDVLSEIMNSFAKSIEVMGDGCLNDEHL
                     EELGGILKAKLEGHPKNQELRQVKRQEENYDQQVEMSLQDEDECDVYILTKVSDILHS
                     LFSTYKEKILPWFEQLLPLIVNLICSSRPWPDRQWGLCIFDDIIEHCSPTSFKYVEYF
                     RWPMLLNMRDNNPEVRQAAAYGLGVMAQFGGDDYRSLCSEAVPLLVKVIKCANSKTKK
                     NVIATENCISAIGKILKFKPNCVNVDEVLPHWLSWLPLHEDKEEAIQTLSFLCDLIES
                     NHPVVIGPNNSNLPKIISIIAEGKINETINYEDPCAKRLANVVRQVQTSEDLWLECVS
                     QLDDEQQEALQELLNFA"
     gene            complement(join(<9894. .10679,11795. .11935))
                     /locus_tag="RP11-218I7.1-002"
     mRNA            complement(join(<9894. .10679,11795. .11935))
                     /locus_tag="RP11-218I7.1-002"
                     /note="match: ESTs: AA043342.1 AL527785.1 BI918575.1"
     CDS             complement(join(<9894. .10679,11795. .11809))
                     /locus_tag="RP11-218I7.1-002"
                     /standard_name="OTTHUMP00000021040"
                     /codon_start=1
                     /protein_id="CAI12714.1"
                     /db_xref="GI:55957961"
                     /db_xref="GOA:Q5T7X3"
                     /db_xref="InterPro:IPR000357"
                     /db_xref="UniProtKB/TrEMBL:Q5T7X3"
                     /translation="MKISQDPHPRVRAAACTTLGQMATDFAPNFQKKFHETVIAALLR
                     TMENQGNQRVQSHAASALIIFIEDCPKSLLVLYVDSMVKNLHSVLVIKLQELIRNGTK
                     LALEQLVTTIASVADTIEEKFVPYYDIFMPSLKHIVELAVQKELKLLRGKTIECISHI
                     GLAVGKEKFMQDASNVMQLLLKTQSDLNNMEDDDPQTSYMVSAWARMCKILGKDFQQY
                     LPLVIEPLIKTASAKPDVALLDTQDVENMSDDDGWQFVNLGDQQSFGIK"
     misc_feature    35944
                     /note="Clone_right_end: RP11-165O14"
     gene            join(43683. .43831,62688. .63037)
                     /locus_tag="RP11-218I7.2-001"
     mRNA            join(43683. .43831,62688. .63037)
                     /locus_tag="RP11-218I7.2-001"
                     /note="match: ESTs: AI651280.1 AI970391.1 AW243550.1
                     AW590614.1 BF062077.1 BF222880.1"
     misc_feature    68366
                     /note="Clone_left_end: RP11-575C20"
ORIGIN

  Query Match             38.8%;  Score 155;  DB 5;  Length 68465;
  Best Local Similarity   76.0%;  Pred. No. 4.7e-42;
  Matches  209;  Conservative    0;  Mismatches   55;  Indels   11;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||| |||||||||||  |||||||||||||||||| |||
Db       23811 ATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCAGGCGGATCACCTGAGGTCAGGAGT 23870

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || ||||||  |||| |||||||||||||||||||||| |||||||||||||||| ||||
Db       23871 TCAAGACCAGCCTGACCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 23930

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||  |||||||||||| |||||||||||||||||||||||| | |||  || ||
Db       23931 TGGGTGTGGTGGCACACGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGATGGAGAATCT 23990

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAG-----------CAGAGTG 309
                 || | | |   |    ||  ||  || |  ||| || |           ||||| |
Db       23991 CTTGAACCCGGGAGGTAGAGACTGCAGTGAGCCGAGACTGCGCCACTGCATGACAGAGCG 24050

Qy         310 AGACKCCGTCTCAAAAACAACAACAAAAAACAAAA 344
               |||| || ||||||||| || || ||||||   ||
Db       24051 AGACACCATCTCAAAAAAAAAAAAAAAAAAAGGAA 24085


RESULT 50
AC068889
LOCUS       AC068889               73433 bp    DNA     linear   PRI 01-JAN-2003
DEFINITION  Homo sapiens 12 BAC RP11-774I22 (Roswell Park Cancer Institute
            Human BAC Library) complete sequence.
ACCESSION   AC068889
VERSION     AC068889.35  GI:27452896
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 73433)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S.,
            Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P.,
            Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N.,
            Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P.,
            Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M.,
            Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S.,
            Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y.,
            Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U.,
            King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N.,
            Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O.,
            Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X.,
            Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P.,
            Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E.,
            Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S.,
            Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T.,
            Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M.,
            Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 73433)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 73433)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 73433)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JAN-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
```

COMMENT On Jan 1, 2003 this sequence version replaced gi:23264934.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

FEATURES Location/Qualifiers

source	1. .73433
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="12"
	/clone="RP11-774I22"
repeat_region	12. .66
	/rpt_family="(CGGG)n"
misc_feature	75. .2004
	/note="overlaps bases 170628. .172558 of clone AC023509"
	/function="clone overlap"
repeat_region	505. .799
	/rpt_family="L1MC5"
repeat_region	842. .888
	/rpt_family="(TA)n"
repeat_region	915. .1149
	/rpt_family="AluJo"
repeat_region	complement(1166. .1289)
	/rpt_family="AluSp/q"
repeat_region	1308. .1487
	/rpt_family="L1MC4"
repeat_region	1428. .1521
	/rpt_family="L1MC5"
repeat_region	1608. .1623
	/rpt_family="AluJo"
repeat_region	1624. .1923
	/rpt_family="AluSp"
repeat_region	1924. .2066
	/rpt_family="AluJo"
repeat_region	2100. .2162
	/rpt_family="L1MC5"
repeat_region	complement(2163. .2467)
	/rpt_family="AluSg"
repeat_region	2468. .2537
	/rpt_family="L1MC5"
repeat_region	2538. .2846
	/rpt_family="AluSc"
repeat_region	2847. .2962
	/rpt_family="L1MC5"
repeat_region	2968. .3267
	/rpt_family="AluSc"
repeat_region	3269. .3447
	/rpt_family="AluJb"
repeat_region	3457. .3714
	/rpt_family="AluJb"
repeat_region	3872. .4141
	/rpt_family="AluSp"
repeat_region	complement(4142. .4197)
	/rpt_family="Alu"
repeat_region	4538. .4588
	/rpt_family="MIR"
repeat_region	complement(4613. .4854)
	/rpt_family="AluSp"
repeat_region	4904. .5211
	/rpt_family="AluSq"
repeat_region	5220. .5254
	/rpt_family="AT_rich"
STS	5615. .5787
	/standard_name="SHGC-187"
STS	5624. .5762
	/standard_name="D3S4196"
repeat_region	complement(6853. .7147)
	/rpt_family="AluSg"
repeat_region	complement(7303. .7515)
	/rpt_family="AluSg/x"
repeat_region	7589. .7887
	/rpt_family="AluY"
repeat_region	9063. .9327
	/rpt_family="AluSx"
repeat_region	9675. .9955
	/rpt_family="AluSg"
repeat_region	9956. .9978
	/rpt_family="(GAAAA)n"
STS	10267. .10546
	/standard_name="G64247"
repeat_region	10966. .11220
	/rpt_family="L2"
repeat_region	complement(11221. .11509)
	/rpt_family="AluSc"
repeat_region	11524. .11535
	/rpt_family="AluSg/x"
repeat_region	11536. .11839
	/rpt_family="AluSg"
misc_feature	11548
	/note="Position below phred 30 quality."
	/function="low quality"
repeat_region	11840. .12010

Query Match 38.8%; Score 155; DB 5; Length 73433;
Best Local Similarity 75.6%; Pred. No. 4.8e-42;
Matches 208; Conservative 1; Mismatches 56; Indels 10; Gaps 1;

```
Qy        81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
             | |||||||||||||||||||| |||||||||||| ||| |||||| ||||||||| | |
Db     54382 ACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGAGGATCATCTGAGGTCAGGGGT 54441

Qy       141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
             || |||||| || |||||||||||||||||  ||| ||||||||||||||||| || |
Db     54442 TCAAGACCAGCCAGGCCAACATGGTGAAACCCTATCTCTACTAAAAATACAAAAATTATC 54501

Qy       201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
             |||||| ||||||| | |||||| |||||||||||||||| |  || ||||||||||
Db     54502 TGGGCACGGTGGCAGGCGCCTGTAATCCCAGCTACTCAGGAGGCACAGGTTGCAGTGAGC 54561

Qy       261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
             ||||||           | |   ||   ||    | | | ||||   |||:|| |||
Db     54562 CGAGATC----------ATGCCACTGCACTCCAGCCTGGGCGAAGAGCAGGACTCCATCT 54611
```

```
Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGA 355
              |||||| || || ||| || |||||||   | ||||
Db      54612 CAAAAAAAAAAAAAAAGAAAAAAAAAGGAAAAAGA 54646
```

Search completed: July 17, 2006, 21:02:06
Job time : 2733 secs

GenCore version 5.1.9

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 20:02:44 ; Search time 525 Seconds
(without alignments)
5312.189 Million cell updates/sec

Title: SEQ1-47502C
Perfect score: 399.6
Sequence: 1 ccaggtactcagccatgtgc.....tattgcaggagaccaaaaag 400

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : N_Geneseq_8:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*
15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	398	99.6	94400	12	ADP08387	Adp08387 Human gly
	2	158.4	39.6	16163	4	AAI62620	Aai62620 Human bre
c	3	158.4	39.6	16163	4	AAK83901	Aak83901 Human imm
c	4	158.4	39.6	16163	4	AAK87161	Aak87161 Human imm
	5	158.4	39.6	16163	4	AAL03817	Aal03817 Human rep
	6	158.4	39.6	96593	9	ADA02750	Ada02750 Human IRF
	7	158.4	39.6	96593	10	ADB72488	Adb72488 Human IRF
	8	158.4	39.6	96593	10	ADC85230	Adc85230 Human Irf
	9	158.4	39.6	96593	12	ADM74345	Adm74345 Human car
c	10	155.4	38.9	31898	11	ACN44354	Acn44354 Human gen
	11	154.4	38.6	34796	13	ACN37240	Acn37240 Human per
c	12	154.2	38.6	122748	6	ABT10719	Abt10719 Human bre
	13	153.6	38.4	884	3	AAC79009	Aac79009 Human sec
	14	153.2	38.3	14784	3	AAA64141	Aaa64141 Nucleotid
	15	152.8	38.2	166942	15	AEE03362	Aee03362 Acute mye
c	16	152.4	38.1	4388	4	AAL04959	Aal04959 Human rep
c	17	152.4	38.1	4388	4	ABL97853	Abl97853 Human tes
c	18	152	38.0	174424	6	ABL68122	Abl68122 Ovary can
c	19	152	38.0	181343	12	ADQ19573	Adq19573 Human sof
c	20	151.4	37.9	169739	6	ABQ88186	Abq88186 Human ost
c	21	150.6	37.7	14282	4	AAK68418	Aak68418 Human imm
c	22	150.6	37.7	32146	4	AAS28363	Aas28363 Genomic s
c	23	150.6	37.7	32146	10	ADG41559	Adg41559 Human res
c	24	150.6	37.7	32146	11	ADI97333	Adi97333 Human res
c	25	150.6	37.7	125515	10	ADL13941	Adl13941 Osteoarth
c	26	150.4	37.6	17538	3	AAZ29204	Aaz29204 Human mye
c	27	150.4	37.6	17538	6	ABN86597	Abn86597 Human mye
c	28	150.4	37.6	88892	12	ADQ97695	Adq97695 Human can
	29	150	37.5	4779	4	AAI61042	Aai61042 Human pol
	30	150	37.5	100998	14	AED18472	Aed18472 Fibrotic
c	31	149.8	37.5	151909	14	AEB96535	Aeb96535 Human CAB
	32	149.8	37.5	196686	11	ACN44170	Acn44170 Human gen
c	33	149.6	37.4	5670	6	ABL65836	Abl65836 Lung canc
c	34	149.6	37.4	5670	8	ACA64883	Aca64883 Human MB-
c	35	149.6	37.4	5670	15	AEF74508	Aef74508 Human pol
c	36	149	37.3	8205	4	AAK90749	Aak90749 Human dig
c	37	149	37.3	8205	4	AAI62948	Aai62948 Human gen
	38	149	37.3	89900	12	ADO79404	Ado79404 DPF3 regi
c	39	149	37.3	174318	15	AEF92655	Aef92655 Human Ras
	40	149	37.3	285300	14	ADX98573	Adx98573 Human D4,
c	41	148.8	37.2	201143	6	ABK83568	Abk83568 Human DNA
c	42	148.6	37.2	44075	9	ADA02774	Ada02774 Human ARH
c	43	148.6	37.2	44075	10	ADB72512	Adb72512 Human ARH
c	44	148.6	37.2	44075	10	ADC85254	Adc85254 Human Arh
c	45	148.6	37.2	44075	12	ADM74369	Adm74369 Human car
	46	148.6	37.2	49745	11	ACN45134	Acn45134 Human gen
	47	148.6	37.2	110000	14	AEE04996_2	Continuation (3 of
	48	148.6	37.2	133893	9	AAD54538	Aad54538 Human pho
	49	148.4	37.1	122557	12	ADH76849	Adh76849 Melanin-c
	50	148	37.0	9469	4	AAK79514	Aak79514 Human imm
	51	148	37.0	143973	11	ACN44142	Acn44142 Human gen
	52	147.8	37.0	286	4	AAK77171	Aak77171 Human imm
	53	147.8	37.0	2030	15	AEE83631	Aee83631 Human cDN
	54	147.8	37.0	270150	11	ADP65796	Adp65796 Human 16p
	55	147.6	36.9	11820	3	AAA95944	Aaa95944 Human KLK
	56	147.6	36.9	11820	12	ADN10929	Adn10929 Human kal
	57	147.6	36.9	17804	13	ADS36476	Ads36476 Human aut
	58	147.6	36.9	78268	11	ACN44342	Acn44342 Human gen
c	59	147.4	36.9	6092	12	ADQ83572	Adq83572 Human tum
	60	147.4	36.9	90798	12	ADP68859	Adp68859 Human chr
c	61	147.4	36.9	98634	11	ACN45094	Acn45094 Human gen
	62	147.2	36.8	47219	13	ADR16283	Adr16283 Human Cay
	63	147.2	36.8	59884	13	ADR16284	Adr16284 Human Cay
c	64	147.2	36.8	110000	13	ABD32911_5	Continuation (6 of
c	65	147.2	36.8	110000	14	ADZ46976_0	Adz46976 Human pho
c	66	147.2	36.8	209613	14	ADY25743	Ady25743 Uridine p
c	67	147	36.8	2909	11	ADM02428	Adm02428 Human cDN
c	68	147	36.8	2909	14	AEC85358	Aec85358 Human cDN
c	69	147	36.8	3623	12	ADQ63217	Adq63217 Novel hum
c	70	147	36.8	3751	13	ADR08282	Adr08282 Full leng
c	71	147	36.8	6899	10	ADG42143	Adg42143 Human bra
	72	147	36.8	30037	5	ABA20038	Aba20038 Human ner
c	73	147	36.8	57013	11	ACN45046	Acn45046 Human gen
	74	147	36.8	238417	13	ABD32868	Abd32868 Human can
	75	146.8	36.7	2061	2	AAX00614	Aax00614 Human sec
	76	146.8	36.7	247509	10	ADL13684	Adl13684 Osteoarth
c	77	146.6	36.7	24318	13	ABD33393	Abd33393 Human can
	78	146.6	36.7	51961	10	ADC20666	Adc20666 Human sec
	79	146.6	36.7	51961	10	ABT16905	Abt16905 Human sec
	80	146.6	36.7	51961	10	ABZ67488	Abz67488 Human sec
c	81	146.6	36.7	260160	12	ADQ20017	Adq20017 Human sof
c	82	146.2	36.6	10859	4	AAL04720	Aal04720 Human rep
c	83	146.2	36.6	10859	4	ABL97627	Abl97627 Human tes
c	84	146.2	36.6	163218	15	AEF92730	Aef92730 Human chr
c	85	146	36.5	444	5	ABV04721	Abv04721 Human pro
c	86	146	36.5	190000	10	ADL13752	Adl13752 Osteoarth
	87	145.8	36.5	32249	4	AAL05336	Aal05336 Human rep
	88	145.8	36.5	32249	4	ABL98205	Abl98205 Human tes
	89	145.8	36.5	115756	8	ACD13448	Acd13448 Human DNA
	90	145.6	36.4	5161	4	AAS36060	Aas36060 Human car
	91	145.6	36.4	5161	10	ADE46754	Ade46754 Human car

	92	145.6	36.4	5161	13	ADJ08172 Adj08172 Human car
c	93	145.6	36.4	94400	12	ADP08387 Adp08387 Human gly
	94	145.6	36.4	106130	12	ADQ97599 Adq97599 Human can
c	95	145.4	36.4	2566	4	AAH76802 Aah76802 Human thy
c	96	145.4	36.4	86080	6	ABQ88164 Abq88164 Human ost
c	97	145.4	36.4	86080	6	ABK83561 Abk83561 Human cDN
c	98	145.4	36.4	86080	10	ADD71054 Add71054 Human pro
c	99	145.4	36.4	86080	12	ADQ18878 Adq18878 Human sof
c	100	145.4	36.4	86080	14	AED18140 Aed18140 Fibrotic
c	101	145.2	36.3	21535	4	AAS29834 Aas29834 Human cyt
c	102	145.2	36.3	21535	4	AAS35081 Aas35081 DNA #31 e
c	103	145.2	36.3	21535	10	ADC46523 Adc46523 Human neo
c	104	145.2	36.3	81001	4	AAF30035 Aaf30035 Human apo
	105	145.2	36.3	188971	12	ADL08108 Adl08108 Human gen
c	106	145.2	36.3	220895	6	ABK84798 Abk84798 Human cDN
c	107	145.2	36.3	220895	13	ADR52737 Adr52737 Drug ther
c	108	145	36.3	3324	13	ADR08168 Adr08168 Full leng
c	109	145	36.3	20522	4	AAL37409 Aal37409 Human mus
c	110	145	36.3	20522	8	ABX60397 Abx60397 cDNA enco
c	111	145	36.3	20522	12	ADJ31147 Adj31147 Human mus
c	112	145	36.3	86000	12	ADP68568 Adp68568 Human PPA
	113	145	36.3	92638	6	ABQ88096 Abq88096 Human ost
	114	145	36.3	110000	13	ABD32780_3 Continuation (4 of
c	115	144.8	36.2	66566	3	AAA53450 Aaa53450 Human thi
c	116	144.8	36.2	295772	12	ADQ97433 Adq97433 Human can
c	117	144.6	36.2	2744	4	AAK79904 Aak79904 Human imm
c	118	144.6	36.2	3670	12	ADQ63218 Adq63218 Novel hum
c	119	144.6	36.2	6512	12	ADN05238 Adn05238 Antipsori
c	120	144.6	36.2	6512	14	ADY18445 Ady18445 DNA encod
c	121	144.6	36.2	48037	4	AAK84729 Aak84729 Human imm
c	122	144.6	36.2	48037	4	AAK85983 Aak85983 Human imm
c	123	144.6	36.2	48045	4	AAK84730 Aak84730 Human imm
c	124	144.6	36.2	48045	4	AAK85984 Aak85984 Human imm
	125	144.6	36.2	59725	11	ACN44390 Acn44390 Human gen
c	126	144.6	36.2	84476	12	ADQ97028 Adq97028 Human can
c	127	144.6	36.2	110000	14	ADZ45062_03 Continuation (4 of
c	128	144.6	36.2	110000	14	ADZ45062_04 Continuation (5 of
c	129	144.6	36.2	174472	14	ADZ13139 Adz13139 Human can
c	130	144.6	36.2	174703	11	ACN44738 Acn44738 Human gen
	131	144.6	36.2	207433	5	ABZ72040 Abz72040 Gene 216
	132	144.6	36.2	207433	8	ABX74891 Abx74891 BAC1098L2
	133	144.6	36.2	207433	12	ADJ36614 Adj36614 Bacterial
	134	144.6	36.2	207433	12	ADL81193 Adl81193 BAC1098L2
c	135	144.6	36.2	207600	14	AED89391 Aed89391 Human bre
	136	144.4	36.1	314	12	ADL70204 Adl70204 Human gon
	137	144.4	36.1	1965	6	ABK34980 Abk34980 Human cDN
	138	144.4	36.1	10800	4	AAK83329 Aak83329 Human imm
c	139	144.4	36.1	43064	8	ACC80906 Acc80906 Human CYP
	140	144.4	36.1	65454	12	ADN01773 Adn01773 Human hun
	141	144.4	36.1	86764	12	ADQ97623 Adq97623 Human can
c	142	144.4	36.1	86950	14	ADX98572 Adx98572 Human HT0
	143	144.4	36.1	227968	6	ABK83497 Abk83497 Human cDN
	144	144.4	36.1	227968	12	ADQ18538 Adq18538 Human sof
	145	144.4	36.1	227968	14	AED17856 Aed17856 Fibrotic
c	146	144.2	36.1	4562	4	AAK74521 Aak74521 Human imm
c	147	144.2	36.1	167343	6	ABL64403 Abl64403 Stomach c
c	148	144.2	36.1	167343	6	ABL67239 Abl67239 Thyroid c
c	149	144.2	36.1	227968	6	ABK83497 Abk83497 Human cDN
c	150	144.2	36.1	227968	12	ADQ18538 Adq18538 Human sof

ALIGNMENTS

```
RESULT 1
ADP08387
ID   ADP08387 standard; DNA; 94400 BP.
XX
AC   ADP08387;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human glycoprotein VI (platelet) (GP6;GPIV;GPVI) genomic DNA.
XX
KW   breast cancer; cytostatic; gene therapy; human; platelet glycoprotein VI;
KW   GP6; GPIV; GPVI; chromosome 19q13.4; ds; gene; SNP;
KW   single nucleotide polymorphism.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       185
FT                   /*tag= a
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       237
FT                   /*tag= b
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       641
FT                   /*tag= c
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       719
FT                   /*tag= d
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       990
FT                   /*tag= e
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       2908
FT                   /*tag= f
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       3140
FT                   /*tag= g
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       3880
FT                   /*tag= h
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       4494
FT                   /*tag= i
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       5107
FT                   /*tag= j
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       5220
FT                   /*tag= k
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       6031
FT                   /*tag= l
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       8670
FT                   /*tag= m
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       13794
FT                   /*tag= n
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       16356
FT                   /*tag= o
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       17164
FT                   /*tag= p
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       17264
FT                   /*tag= q
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       20537
FT                   /*tag= r
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       20637
FT                   /*tag= s
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       20900
FT                   /*tag= t
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       21155
FT                   /*tag= u
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       21795
FT                   /*tag= v
```

```
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       21931
FT                   /*tag= w
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       22167
FT                   /*tag= x
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       22656
FT                   /*tag= y
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       23108
FT                   /*tag= z
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       23404
FT                   /*tag= aa
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       24287
FT                   /*tag= ab
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       24480
FT                   /*tag= ac
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       24592
FT                   /*tag= ad
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       24878
FT                   /*tag= ae
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       26370
FT                   /*tag= af
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       27056
FT                   /*tag= ag
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       27874
FT                   /*tag= ah
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       31248
FT                   /*tag= ai
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       31458
FT                   /*tag= aj
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       31553
FT                   /*tag= ak
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       31637
FT                   /*tag= al
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       31668
FT                   /*tag= am
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       31752
FT                   /*tag= an
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       37643
FT                   /*tag= ao
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       43941
FT                   /*tag= ap
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       44134
FT                   /*tag= aq
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       44329
FT                   /*tag= ar
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       44343
FT                   /*tag= as
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       44362
FT                   /*tag= at
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       44818
FT                   /*tag= au
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       44917
FT                   /*tag= av
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       45215
FT                   /*tag= aw
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       45666
FT                   /*tag= ax
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       45680
FT                   /*tag= ay
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       46402
FT                   /*tag= az
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       46510
FT                   /*tag= ba
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       46554
FT                   /*tag= bb
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       46823
FT                   /*tag= bc
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       47714
FT                   /*tag= bd
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       48963
FT                   /*tag= be
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       49157
FT                   /*tag= bf
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       49254
FT                   /*tag= bg
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       49257
FT                   /*tag= bh
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       49356
FT                   /*tag= bi
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       55202
FT                   /*tag= bj

  Query Match             99.6%;  Score 398;  DB 12;  Length 94400;
  Best Local Similarity   99.8%;  Pred. No. 1.7e-100;
  Matches  399;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCAGGTACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47401 CCAGGTACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGG 47460

Qy         61 CCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCG 120
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      47461 CCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCG 47520

Qy        121 GATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47521 GATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTA 47580

Qy        181 CTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47581 CTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGG 47640

Qy        241 AGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47641 AGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTG 47700

Qy        301 AGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 360
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47701 AGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 47760

Qy        361 TCCATCTGCGGTTCCCAGACTATTGCAGGAGACCAAAAAG 400
                ||||||||||||||||||||||||||||||||||||||||
Db      47761 TCCATCTGCGGTTCCCAGACTATTGCAGGAGACCAAAAAG 47800


RESULT 2
AAI62620
ID   AAI62620 standard; DNA; 16163 BP.
XX
AC   AAI62620;
XX
DT   19-OCT-2001  (first entry)
XX
DE   Human breast or ovarian antigen genomic DNA SEQ ID NO: 270.
XX
KW   Human; breast antigen; ovarian antigen; cancer; metastasis; gene therapy;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200155324-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001344.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241221P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
```

```
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-488785/53.
XX
PT   New isolated nucleic acids and polypeptides, useful for diagnosing,
PT   treating and/or preventing human diseases and disorders.
XX
PS   Disclosure; SEQ ID NO 270; 520pp + Sequence Listing; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of ovarian and breast antigens. These are shown in AAI62467-
CC   AAI62572 and AAM42240-AAM42345. The sequences can be used in the
CC   diagnosis, prevention and treatment of breast and ovarian cancers, and
CC   their metastases. The present sequence is a genomic sequence of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 16163 BP; 5062 A; 3448 C; 3239 G; 4414 T; 0 U; 0 Other;

  Query Match            39.6%;  Score 158.4;  DB 4;  Length 16163;
  Best Local Similarity  73.9%;  Pred. No. 1e-33;
  Matches  201;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy         76 GCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCA 135
              ||  | ||||||||||||||||||  |||||||| ||| ||| |||||||||| ||||
Db       8172 GGTTCACGCCTGTAATCCCAGCACTCTGGGAGGCCGAGGCGGGTGGATCACCTGATGTCA 8231

Qy        136 AGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAA 195
              ||| | |||| ||  |||||||||||||||||||||||||| |||||||  |||||||||
Db       8232 GGAGTTTGAGATCAGTCTGGCCAACATGGTGAAACCCCGTCTCTACTAAATATACAAAAA 8291

Qy        196 ATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG 255
              |||||||  |||||| |||| | |||| |||||||||||| |||| ||||| ||||||
Db       8292 CCAGCTGGGTGTGGTGGTACACGCTTGTAATCCCAGCTACTCTGGAGGCGGAGGTTGCAG 8351

Qy        256 TGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKC 315
              ||||| || |  |||      | |   | |   | | ||||||| |||          |
Db       8352 TGAGCCGAAACTGCACCACTACACTCCAGCCTGGGTGACAGAGGGAGACTCTGTCTCCCA 8411

Qy        316 CGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
              |    ||||| || || ||||||| |||||||
Db       8412 GGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 8443


RESULT 3
AAK83901/c
ID   AAK83901 standard; DNA; 16163 BP.
XX
AC   AAK83901;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38713.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
```

PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 38713; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 16163 BP; 4414 A; 3239 C; 3448 G; 5062 T; 0 U; 0 Other;

Query Match 39.6%; Score 158.4; DB 4; Length 16163;
Best Local Similarity 73.9%; Pred. No. 1e-33;
Matches 201; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```
Qy       76 GCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCA 135
           | |  | ||||||||||||||||||  |||||||| ||| ||| |||||||||| ||||
Db     7992 GGTTCACGCCTGTAATCCCAGCACTCTGGGAGGCCGAGGCGGGTGGATCACCTGATGTCA 7933

Qy      136 AGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAA 195
            ||| | |||| ||  ||||||||||||||||||||||||||||| ||||||| |||||||||
Db     7932 GGAGTTTGAGATCAGTCTGGCCAACATGGTGAAACCCCGTCTCTACTAAATATACAAAAA 7873

Qy      196 ATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG 255
              |||||||  |||||| |||| | |||| |||||||||||| |||| ||||| ||||||
Db     7872 CCAGCTGGGTGTGGTGGTACACGCTTGTAATCCCAGCTACTCTGGAGGCGGAGGTTGCAG 7813

Qy      256 TGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKC 315
            ||||| || |  |||     |  |   | |   | | |||||| |||           |
Db     7812 TGAGCCGAAACTGCACCACTACACTCCAGCCTGGGTGACAGAGGGAGACTCTGTCTCCCA 7753

Qy      316 CGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
              |    ||||| || || |||||| |||||||
Db     7752 GGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 7721
```

RESULT 4
AAK87161/c
ID AAK87161 standard; DNA; 16163 BP.
XX
AC AAK87161;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:41973.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

```
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS   Disclosure; SEQ ID NO 41973; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting the
CC   nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention
XX
SQ   Sequence 16163 BP; 4414 A; 3239 C; 3448 G; 5062 T; 0 U; 0 Other;

  Query Match             39.6%;  Score 158.4;  DB 4;  Length 16163;
  Best Local Similarity   73.9%;  Pred. No. 1e-33;
  Matches  201;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy         76 GCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCA 135
              | |  | ||||||||||||||||||  |||||||| ||| ||| |||||||||||| ||||
Db       7992 GGTTCACGCCTGTAATCCCAGCACTCTGGGAGGCCGAGGCGGGTGGATCACCTGATGTCA 7933

Qy        136 AGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAA 195
               ||| | |||| ||  |||||||||||||||||||||||||| ||||||| |||||||||
Db       7932 GGAGTTTGAGATCAGTCTGGCCAACATGGTGAAACCCCGTCTCTACTAAATATACAAAAA 7873

Qy        196 ATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG 255
               |||||||  |||||| |||| | |||| |||||||||||| |||| ||||| ||||||
Db       7872 CCAGCTGGGTGTGGTGGTACACGCTTGTAATCCCAGCTACTCTGGAGGCGGAGGTTGCAG 7813

Qy        256 TGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKC 315
              |||||  || |  |||      |  |    | |    | | |||||| |||          |
Db       7812 TGAGCCGAAACTGCACCACTACACTCCAGCCTGGGTGACAGAGGGAGACTCTGTCTCCCA 7753

Qy        316 CGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
               |    ||||| || || ||||||| |||||||
Db       7752 GGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 7721


RESULT 5
AAL03817
ID   AAL03817 standard; DNA; 16163 BP.
XX
AC   AAL03817;
XX
DT   21-NOV-2001 (first entry)
XX
DE   Human reproductive system related antigen DNA SEQ ID NO: 6505.
XX
KW   Human; reproductive system related antigen; reproductive system disorder;
KW   cancer; gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155320-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001339.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
```

PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465570/50.
XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen is
PT used in preventing, treating or ameliorating a medical condition.
XX
PS Disclosure; SEQ ID NO 6505; 1297pp + Sequence Listing; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention
XX
SQ Sequence 16163 BP; 5062 A; 3448 C; 3239 G; 4414 T; 0 U; 0 Other;

Query Match 39.6%; Score 158.4; DB 4; Length 16163;
Best Local Similarity 73.9%; Pred. No. 1e-33;
Matches 201; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 76 GCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCA 135
Db 8172 GGTTCACGCCTGTAATCCCAGCACTCTGGGAGGCCGAGGCGGGTGGATCACCTGATGTCA 8231

Qy 136 AGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAA 195
Db 8232 GGAGTTTGAGATCAGTCTGGCCAACATGGTGAAACCCCGTCTCTACTAAATATACAAAAA 8291

Qy 196 ATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG 255
Db 8292 CCAGCTGGGTGTGGTGGTACACGCTTGTAATCCCAGCTACTCTGGAGGCGGAGGTTGCAG 8351

Qy 256 TGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKC 315
Db 8352 TGAGCCGAAACTGCACCACTACACTCCAGCCTGGGTGACAGAGGGAGACTCTGTCTCCCA 8411

Qy 316 CGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
Db 8412 GGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 8443

RESULT 6
ADA02750
ID ADA02750 standard; DNA; 96593 BP.
XX
AC ADA02750;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human IRF2 carcinoma associated gene, SEQ ID NO:1268.
XX
KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003057146-A2.
XX
PD 17-JUL-2003.
XX
PF 26-DEC-2002; 2002WO-US041414.
XX
PR 26-DEC-2001; 2001US-00035832.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-587068/55.
XX
PT New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.
XX

```
PS   Claim 1; SEQ ID NO 1268; 245pp; English.
XX
CC   The invention relates to recombinant carcinoma associated (CA) nucleic
CC   acid sequences from mouse and human (ADA01482-ADA03094), and to
CC   recombinant carcinoma associated proteins (CAP) encoded by them. The
CC   invention also encompasses expression vectors and host cells comprising a
CC   CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC   binds to the protein, and a biochip comprising CA nucleic acid or
CC   fragments thereof. The sequences of the invention were identified using
CC   oncogenic retroviruses, which insert into the genome of the host organism
CC   at random. Many of these do not carry transduced host oncogenes or
CC   pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC   direct consequence of the effects of proviral integration into host
CC   protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC   carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC   leukaemia) or a propensity to carcinoma by determination of the sequence
CC   of a CA gene, or by determination of CA gene expression in particular
CC   tissues. CA nucleic acids, proteins and antibodies are also useful as
CC   therapeutic agents and in screening and evaluating drug candidates. The
CC   present sequence represents a specifically claimed human CA nucleic acid
CC   sequence of the invention. Note: The complete sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 96593 BP; 25039 A; 20903 C; 22979 G; 27672 T; 0 U; 0 Other;

  Query Match             39.6%;  Score 158.4;  DB 9;  Length 96593;
  Best Local Similarity   75.0%;  Pred. No. 1.8e-33;
  Matches  216;  Conservative    1;  Mismatches   57;  Indels   14;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||||||||||||||| |||||||| ||| ||||||||  |||||||||| |||
Db       18217 ACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATTTCCTGAGGTCAGGAGT 18276

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || ||| || ||||||||||||||||||||||||||| |||||||||||||||| ||||
Db       18277 TCAAGATCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 18336

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||||||||||   |||||||||||||||||||| |||| | |||   | ||
Db       18337 TGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCG 18396

Qy         261 TGAGATCGCAG--------------AGTGAGCCGAAATCACAGATCACAGAGTGAGCAGA 306
                 || | |||               |||||||||| ||| ||       | ||||  ||
Db       18397 CTTGAACCCAGGCGGCGGAGGTTGCAGTGAGCCGAGATCGCACCAGCTTGGGTGACAAGC 18456

Qy         307 GTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAG 354
               |||||||: | ||||||||| ||  | |||| | |||| ||||| | |
Db       18457 GTGAGACTTCATCTCAAAAAAAAACAAAAAACAAAAAACAACCAGATG 18504


RESULT 7
ADB72488
ID   ADB72488 standard; DNA; 96593 BP.
XX
AC   ADB72488;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human IRF2 gene.
XX
KW   human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW   cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS   Homo sapiens.
XX
PN   WO2003008583-A2.
XX
PD   30-JAN-2003.
XX
PF   26-DEC-2001; 2001WO-US051291.
XX
PR   02-MAR-2001; 2001US-00798586.
PR   23-OCT-2001; 2001US-00004113.
PR   08-NOV-2001; 2001US-00052482.
PR   30-NOV-2001; 2001US-00997722.
PR   20-DEC-2001; 2001US-00034650.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW,  Engelhard EK;
XX
DR   WPI; 2003-239337/23.
XX
PT   New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT   cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
PS   Claim 1; SEQ ID NO 316; 2304pp; English.
XX
CC   The invention relates to a novel recombinant nucleic acid comprising a
CC   nucleotide sequence selected from any of the 660 sequences fully defined
CC   in the specification. A polynucleotide of the invention has cytostatic
CC   activity, and may have a use in gene therapy, or in a vaccine. The
CC   recombinant nucleic acids and polypeptides are useful for treating
CC   carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC   sarcomas. The present sequence represents a human gene of the invention.
XX
SQ   Sequence 96593 BP; 25039 A; 20903 C; 22979 G; 27672 T; 0 U; 0 Other;

  Query Match             39.6%;  Score 158.4;  DB 10;  Length 96593;
  Best Local Similarity   75.0%;  Pred. No. 1.8e-33;
  Matches  216;  Conservative    1;  Mismatches   57;  Indels   14;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||||||||||||||| |||||||| ||| ||||||||  |||||||||| |||
Db       18217 ACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATTTCCTGAGGTCAGGAGT 18276

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || ||| || ||||||||||||||||||||||||||| |||||||||||||||| ||||
Db       18277 TCAAGATCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 18336

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||||||||||   |||||||||||||||||||| |||| | |||   | ||
Db       18337 TGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCG 18396

Qy         261 TGAGATCGCAG--------------AGTGAGCCGAAATCACAGATCACAGAGTGAGCAGA 306
                 || | |||               |||||||||| ||| ||       | ||||  ||
Db       18397 CTTGAACCCAGGCGGCGGAGGTTGCAGTGAGCCGAGATCGCACCAGCTTGGGTGACAAGC 18456

Qy         307 GTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAG 354
               |||||||: | ||||||||| ||  | |||| | |||| ||||| | |
Db       18457 GTGAGACTTCATCTCAAAAAAAAACAAAAAACAAAAAACAACCAGATG 18504


RESULT 8
ADC85230
ID   ADC85230 standard; DNA; 96593 BP.
XX
AC   ADC85230;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Human Irf2 genomic sequence.
XX
KW   Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
KW   secreted; transmembrane; intracellular; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003045230-A2.
XX
PD   05-JUN-2003.
XX
PF   02-DEC-2002; 2002WO-US038582.
XX
```

```
PR   30-NOV-2001; 2001US-00997722.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW,  Engelhard EK;
XX
DR   WPI; 2003-513603/48.
XX
PT   New recombinant nucleic acid comprising a nucleotide sequence of any of
PT   the carcinoma-associated (CA) genes, useful for screening for drug
PT   candidates for diagnosing or treating carcinomas.
XX
PS   Claim 1; SEQ ID NO 16; 983pp; English.
XX
CC   The invention relates to a recombinant nucleic acid comprising a
CC   nucleotide sequence selected from any of the fully defined carcinoma-
CC   associated (CA) genes from the 50 tables given in the specification. The
CC   CA proteins are secreted, transmembrane or intracellular proteins. The
CC   recombinant nucleic acids are useful for screening for drug candidates
CC   for diagnosing or treating carcinomas. Sequences given in ADC85215-
CC   ADC85514 represent CA genes of the invention.
XX
SQ   Sequence 96593 BP; 25039 A; 20903 C; 22979 G; 27672 T; 0 U; 0 Other;

  Query Match             39.6%;  Score 158.4;  DB 10;  Length 96593;
  Best Local Similarity   75.0%;  Pred. No. 1.8e-33;
  Matches  216;  Conservative    1;  Mismatches   57;  Indels   14;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| |||||||| ||| ||||||||  |||||||||| |||
Db      18217 ACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATTTCCTGAGGTCAGGAGT 18276

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              || ||| || |||||||||||||||||||||||||||| |||||||||||||||||| ||||
Db      18277 TCAAGATCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCC 18336

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||||||||   |||||||||||||||||||||| |||| | |||   | ||
Db      18337 TGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCG 18396

Qy        261 TGAGATCGCAG--------------AGTGAGCCGAAATCACAGATCACAGAGTGAGCAGA 306
                 || | |||               |||||||||| ||| ||       | ||||  ||
Db      18397 CTTGAACCCAGGCGGCGGAGGTTGCAGTGAGCCGAGATCGCACCAGCTTGGGTGACAAGC 18456

Qy        307 GTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAG 354
              |||||||: | |||||||||| ||  | |||| | |||| ||||| | |
Db      18457 GTGAGACTTCATCTCAAAAAAAAACAAAAAACAAAAAACAACCAGATG 18504


RESULT 9
ADM74345
ID   ADM74345 standard; DNA; 96593 BP.
XX
AC   ADM74345;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Human carcinoma associated (CA) nucleic acid #7.
XX
KW   Human; carcinoma associated nucleic acid; CA nucleic acid; gene; ds;
KW   carcinoma associated protein; CAP; carcinoma; leukaemia; lymphoma;
KW   cytostatic.
XX
OS   Homo sapiens.
XX
PN   US2004072154-A1.
XX
PD   15-APR-2004.
XX
PF   30-NOV-2001; 2001US-00997722.
XX
PR   22-DEC-2000; 2000US-00747377.
PR   02-MAR-2001; 2001US-00798586.
XX
PA   (MORR/) MORRIS D W.
PA   (ENGE/) ENGELHARD E K.
XX
PI   Morris DW,  Engelhard EK;
XX
DR   WPI; 2004-328562/30.
XX
PT   New carcinoma associated gene or protein, useful for preparing a
PT   composition for diagnosing or treating carcinoma e.g., leukemia or
PT   lymphoma.
XX
PS   Claim 1; SEQ ID NO 16; 29pp; English.
XX
CC   The invention relates to new recombinant nucleic acids. The invention
CC   also relates to a host cell comprising a recombinant nucleic acid or
CC   expression vector, an expression vector comprising a recombinant nucleic
CC   acid, a recombinant protein, a method of screening for drug candidates, a
CC   method of screening for a bioactive agent capable of binding to a
CC   carcinoma associated protein (CAP) encoded by a nucleotide sequence, a
CC   method of screening for a bioactive agent capable of modulating the
CC   activity of a CAP, a method of evaluating the effect of a candidate
CC   carcinoma drug, a method of diagnosing carcinoma, a method for inhibiting
CC   the activity of a CAP, a method of treating carcinomas, a method of
CC   neutralising the effect of a CAP and a method of diagnosing carcinoma or
CC   propensity to carcinoma. A method of evaluating the effect of a candidate
CC   carcinoma drug comprises administering the drug to a patient, removing a
CC   cell sample from the patient and determining alterations in the
CC   expression or activation of a gene comprising the nucleotide sequence. A
CC   method of diagnosing carcinoma comprises determining the expression of
CC   one or more genes comprising the nucleic acid sequence in a first tissue
CC   type of a first individual and comparing the expression of the gene from
CC   a second normal tissue type from the first individual or a second
CC   unaffected individual, where a difference in the expression indicates
CC   that the first individual has carcinoma. A method of inhibiting the
CC   activity of a CAP comprises binding an inhibitor to the CAP. Treating
CC   carcinomas comprises administering to a patient an inhibitor of CAP.
CC   Neutralising the effect of a CAP comprises contacting an agent specific
CC   for the CAP. The polypeptide specifically binds to the protein encoded by
CC   the nucleic acid. It comprises an antibody that specifically binds to the
CC   protein encoded by the nucleic acid. The nucleic acids are useful for
CC   preparing a composition for diagnosing or treating carcinoma e.g.,
CC   leukaemia or lymphoma. This sequence represents a human carcinoma
CC   associated (CA) nucleic acid of the invention. Note: The sequence data
CC   for this patent did not form part of the printed specification but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 96593 BP; 25039 A; 20903 C; 22979 G; 27672 T; 0 U; 0 Other;

  Query Match             39.6%;  Score 158.4;  DB 12;  Length 96593;
  Best Local Similarity   75.0%;  Pred. No. 1.8e-33;
  Matches  216;  Conservative    1;  Mismatches   57;  Indels   14;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| |||||||| ||| ||||||||  |||||||||| |||
Db      18217 ACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATTTCCTGAGGTCAGGAGT 18276

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              || ||| || |||||||||||||||||||||||||||| |||||||||||||||||| ||||
Db      18277 TCAAGATCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCC 18336

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||||||||   |||||||||||||||||||||| |||| | |||   | ||
Db      18337 TGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCG 18396

Qy        261 TGAGATCGCAG--------------AGTGAGCCGAAATCACAGATCACAGAGTGAGCAGA 306
                 || | |||               |||||||||| ||| ||       | ||||  ||
Db      18397 CTTGAACCCAGGCGGCGGAGGTTGCAGTGAGCCGAGATCGCACCAGCTTGGGTGACAAGC 18456

Qy        307 GTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAG 354
              |||||||: | |||||||||| ||  | |||| | |||| ||||| | |
Db      18457 GTGAGACTTCATCTCAAAAAAAAACAAAAAACAAAAAACAACCAGATG 18504
```

```
RESULT 10
ACN44354/c
ID   ACN44354 standard; DNA; 31898 BP.
XX
AC   ACN44354;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human genomic sequence hCG17346.
XX
KW   Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003073826-A2.
XX
PD   12-SEP-2003.
XX
PF   28-FEB-2003; 2003WO-US006235.
XX
PR   01-MAR-2002; 2002US-00087192.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW;
XX
DR   WPI; 2003-328604/31.
XX
PT   Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT   comprises a nucleotide sequence.
XX
PS   Claim 1; SEQ ID NO 760; 0pp; English.
XX
CC   The present invention relates to novel DNA and protein sequences which
CC   are associated with carcinomas. The sequences are useful for: (i) for
CC   screening drug candidates; (ii) for screening of bioactive agent capable
CC   of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC   a bioactive agent capable of modulating the activity of CAP; (iv) for
CC   evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC   carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC   carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC   (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC   determining Carcinoma Associated (CA) gene copy number. In addition, the
CC   CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC   carcinoma including lymphoma. The present sequence is one such CA coding
CC   sequence. Note: This patent is an equivalent to basic patent
CC   US2002182586A1, for which no sequence data was published
XX
SQ   Sequence 31898 BP; 7315 A; 9118 C; 8770 G; 6695 T; 0 U; 0 Other;

  Query Match             38.9%;  Score 155.4;  DB 11;  Length 31898;
  Best Local Similarity   72.6%;  Pred. No. 8.8e-33;
  Matches  217;  Conservative    1;  Mismatches   72;  Indels    9;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||| |||||||||||||| || ||||||||||||||||||||| ||||| || ||
Db      30243 ACGCCTATAATCCCAGCACTTTGGAAGGCCAAGGTGGGCGGATCACTTGAGGCCAGGAAT 30184

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| ||||||||| |||||||||||||||| |||||||||||||||| ||||
Db      30183 TCGAGACCAGCCTGGCCAAAATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 30124

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
                ||  ||||||    | |||||||||||| |||||| |||| ||||| ||||||||||
Db      30123 CCGGAGTGGTGGTGGGCGCCTGTAGTCCCAACTACTCGGGAGGCGGAGGTTGCAGTGAGC 30064

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||| | |    |    |  | |   |   |         ||||| |||||:| ||||
Db      30063 CGAGACCACGCCTTTGCACTCCAGCCTGGGCGA---------CAGAGCGAGACTCTGTCT 30013

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGA 379
              ||| ||||| |||||||||  ||| ||    |||   ||| | |  | |  |    ||||
Db      30012 CAACAACAATAACAAAAAAAGAAACAAGTGCAAGGTCTTGCCTTTGAGTTGGGGTCAGA 29954


RESULT 11
ACN37240
ID   ACN37240 standard; DNA; 34796 BP.
XX
AC   ACN37240;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human periodontal disease related gene PLOD SEQ ID NO:150.
XX
KW   periodontal disease; polymorphism; ds; human; gene; SNP;
KW   single nucleotide polymorphism.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    869
FT                   /*tag=  a
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide T,C"
FT   misc_feature    1048
FT                   /*tag=  b
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide T,C"
FT   misc_feature    1103
FT                   /*tag=  c
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide G,A"
FT   misc_feature    5092
FT                   /*tag=  d
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide A,G"
FT   misc_feature    5333
FT                   /*tag=  e
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide C,G"
FT   misc_feature    5795
FT                   /*tag=  f
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide T,C"
FT   misc_feature    6328
FT                   /*tag=  g
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide A,G"
FT   misc_feature    6455
FT                   /*tag=  h
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide A,G"
FT   misc_feature    6489
FT                   /*tag=  i
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide C,G"
FT   misc_feature    6616
FT                   /*tag=  j
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide A,G"
FT   misc_feature    8061
FT                   /*tag=  k
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide G,T"
FT   misc_feature    10001
FT                   /*tag=  l
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide A,G"
FT   misc_feature    12618
FT                   /*tag=  m
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide A,G"
FT   misc_feature    12746
```

```
FT                   /*tag=  n
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide A,G"
FT   misc_feature    15882
FT                   /*tag=  o
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide C,T"
FT   misc_feature    15970
FT                   /*tag=  p
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide C,G"
FT   misc_feature    16961
FT                   /*tag=  q
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide C,T"
FT   misc_feature    17746
FT                   /*tag=  r
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide A,G"
FT   misc_feature    17751
FT                   /*tag=  s
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide C,G"
FT   misc_feature    18004
FT                   /*tag=  t
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide A,G"
FT   misc_feature    22945
FT                   /*tag=  u
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide C,T"
FT   misc_feature    24796
FT                   /*tag=  v
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide C,T"
FT   misc_feature    29622
FT                   /*tag=  w
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide G,A"
FT   misc_feature    29877
FT                   /*tag=  x
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide deletion,CT"
FT   misc_feature    30580
FT                   /*tag=  y
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide G,T"
FT   misc_feature    31500
FT                   /*tag=  z
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide G,A"
FT   misc_feature    32397
FT                   /*tag=  aa
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide C,T"
FT   misc_feature    32406
FT                   /*tag=  ab
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide C,G"
FT   misc_feature    32528
FT                   /*tag=  ac
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide C,G"
FT   misc_feature    33004
FT                   /*tag=  ad
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide C,G"
FT   misc_feature    33056
FT                   /*tag=  ae
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide C,T"
FT   misc_feature    33071
FT                   /*tag=  af
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide G,T"
FT   misc_feature    33127
FT                   /*tag=  ag
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Variable nucleotide A,C"
XX
PN   WO2004042054-A1.
XX
PD   21-MAY-2004.
XX
PF   22-OCT-2003; 2003WO-IB004669.
XX
PR   23-OCT-2002; 2002JP-00308634.
XX
PA   (HUBI-) HUBIT GENOMIX INC.
PA   (KAMO/) KAMOI K.
XX
PI   Kamoi K,  Suzuki A,  Numabe Y,  Ji G,  Muramatsu M,  Baba M;
XX
DR   WPI; 2004-400678/37.
XX
PT   Single nucleotide polymorphisms associated with periodontal disease for
PT   examination and assessment of susceptibility to periodontal disease.
XX
PS   Claim 9; SEQ ID NO 150; 0pp; Japanese.
XX
CC   The invention relates to a novel method for examination of periodontal
CC   disease in which genetic polymorphisms are detected in one or more of 51
CC   genes. The method is useful for examination, diagnosis and assessment of
CC   periodontal disease or risk of periodontal disease and the risk of its
CC   progression to severe, aggressive and chronic periodontal disease. The
CC   present sequence represents a polymorphic gene of the invention
XX
SQ   Sequence 34796 BP; 7631 A; 9152 C; 9368 G; 8645 T; 0 U; 0 Other;

  Query Match           38.6%;  Score 154.4;  DB 13;  Length 34796;
  Best Local Similarity   72.5%;  Pred. No. 1.7e-32;
  Matches  200;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||  |||||| ||||||||||| | | ||||||  || ||||| |||
Db       4654 ATGCCTGTAATCCTGGCACTTTGGGAGGCCAAGGCGTGTGGATCATTTGCGGTCAGGAGT 4713

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||| |||||||||||||| |||||||||| || || |||||||||||||| ||||
Db       4714 TTGAGACCAGCCTGGCCAACATGATGAAACCCCGCCTCTATTAAAAATACAAAAATTAGC 4773

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||||||||      |||||||||||||||||||||||||| ||||| ||||||||||
Db       4774 TGGGCATGGTGGCGAGTGCCTGTAGTCCCAGCTACTCAGGAGGCGGAGGTTGCAGTGAGT 4833

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||||| |           |   | |   | | ||||||||||  |
Db       4834 TGAGATTGTGCCACTGCACTCCAGCTTCGGTGACAGAGTGAGATGCTGTCTCAAAAAAAA 4893

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGAC 356
              ||||  || |  || ||||||| |||  | |  ||
Db       4894 AAAAACAAAAACAAACAAACAAACAAAAAAAACAAC 4929


RESULT 12
ABT10719/c
ID   ABT10719 standard; cDNA; 122748 BP.
XX
AC   ABT10719;
XX
DT   04-DEC-2002  (first entry)
XX
DE   Human breast cancer associated coding sequence SEQ ID NO: 853.
XX
KW   Human; breast specific gene; breast cancer; differential expression;
KW   cytostatic; gene therapy; gene; ss.
```

```
XX
OS   Homo sapiens.
XX
PN   WO200259271-A2.
XX
PD   01-AUG-2002.
XX
PF   25-JAN-2002; 2002WO-US002176.
XX
PR   25-JAN-2001; 2001US-0263757P.
PR   25-APR-2001; 2001US-0286090P.
PR   23-MAY-2001; 2001US-0292517P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Orr MS,  Nation M,  Diggans JC,  Zeng W;
XX
DR   WPI; 2002-674803/72.
XX
PT   Diagnosing breast cancer in a patient comprises detecting the level of
PT   gene expression in cell or tissue samples, where a differential gene
PT   expression is indicative of breast cancer.
XX
PS   Claim 1; SEQ ID NO 853; 260pp + Sequence Listing; English.
XX
CC   The present invention relates to methods of diagnosing breast cancer in a
CC   patient, which comprise detecting the level of expression in a tissue
CC   sample of two or more genes selected from those shown in ABT09867-
CC   ABT11112, where a differential expression of the genes indicates breast
CC   cancer. The methods are useful in diagnosing, treating, detecting the
CC   progression, and in monitoring treatment of breast cancer in patients.
CC   The methods are also useful as a screening tool for agents that modulate
CC   the onset or progression of breast cancer. The breast cancer genes may be
CC   used as diagnostic markers for the prediction or identification of the
CC   malignant state of breast tissue, for confirming the type and progression
CC   of cancer, and for drug screening and assays. The present sequence is a
CC   coding sequence of the invention. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub.published_pct_sequences
XX
SQ   Sequence 122748 BP; 32088 A; 31056 C; 30547 G; 29057 T; 0 U; 0 Other;

  Query Match             38.6%;  Score 154.2;  DB 6;  Length 122748;
  Best Local Similarity   75.1%;  Pred. No. 2.9e-32;
  Matches  208;  Conservative    1;  Mismatches   59;  Indels    9;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||| ||| |||||||| ||| ||| ||||||| | |||||| |||
Db      114310 ATGCCTGTAATCCCAGCGCTTTGGGAGGCCGAGGCGGGTGGATCACTTCAGGTCAGGAGT 114251

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               | ||||||| |||| |||||||||||||||||  || |||||||||||||||| ||||
Db      114250 TTGAGACCAACCTGACCAACATGGTGAAACCCCACCTCTACTAAAAATACAAAAATTAGC 114191

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||| |||||    |||||| |||||||||| |||||| ||||| ||||||||||
Db      114190 TGGGCATGATGGCAGTTGCCTGTAATCCCAGCTACACAGGAGGCGGAGTTTGCAGTGAGT 114131

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               |||||| |    ||    |  | |   |   |          ||||| |||||:||||||
Db      114130 TGAGATTGTGCCGTTGCACTCCAGCCTGGGCGA---------CAGAGCGAGACTCCGTCT 114080

Qy         321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
               |||||| || ||  ||| | || ||||  | |  | |
Db      114079 CAAAAAAAAGAAAGAAAGAAAAGAAAAGAAAAGAAAA 114043


RESULT 13
AAC79009
ID   AAC79009 standard; DNA; 884 BP.
XX
AC   AAC79009;
XX
DT   14-FEB-2001  (first entry)
XX
DE   Human secreted protein gene 13 clone HSREC72.
XX
KW   Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW   antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW   vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW   cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; human; secreted protein; ss.
XX
OS   Homo sapiens.
XX
PN   WO200058358-A1.
XX
PD   05-OCT-2000.
XX
PF   23-MAR-2000; 2000WO-US007725.
XX
PR   26-MAR-1999;   99US-0126602P.
PR   14-JAN-2000; 2000US-0176063P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM,  Komatsoulis G;
XX
DR   WPI; 2000-594640/56.
DR   P-PSDB; AAB44347.
XX
PT   Fourty nine nucleic acid molecules encoding human secreted proteins,
PT   useful in the prevention, treatment and diagnosis of cancer, immune
PT   disorders, cardiovascular disorders and neurological diseases.
XX
PS   Claim 1; Page 322; 367pp; English.
XX
CC   The invention relates to the isolation of genes AAC78997-C79045 encoding
CC   49 human secreted proteins AAB44335-B44382. The genes can be used to
CC   generate fusion proteins by linking to the gene for the human
CC   immunoglobulin G Fc portion (AAC78988) for increasing the stability of
CC   the fusion protein as compared to the human protein only. The genes and
CC   proteins are useful for preventing, ameliorating or treating medical
CC   conditions, e.g. by protein or gene therapy. The genes are isolated from
CC   a range of human tissues disclosed in the specification. The nucleic
CC   acids, proteins, antibodies and (ant)agonists are useful in the
CC   diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC   ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC   marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
CC   immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC   cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC   ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC   infectious diseases such as viral, bacterial, fungal and parasitic
CC   infections
XX
SQ   Sequence 884 BP; 211 A; 214 C; 226 G; 233 T; 0 U; 0 Other;

  Query Match             38.4%;  Score 153.6;  DB 3;  Length 884;
  Best Local Similarity   73.7%;  Pred. No. 9.2e-33;
  Matches  224;  Conservative    1;  Mismatches   70;  Indels    9;  Gaps    2;

Qy          37 AAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAG 96
               | || ||  | | |     |||  ||||||  | ||| |     | ||||||||||| ||
Db      585 ATATTTTGTTTAAATGGAATCATACCAGGCTGGGCACAGTGCTCACGCCTGTAATCCTAG 644

Qy          97 CACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGC 156
               ||||| ||||||||| ||||||| ||||||||||||||  ||| |||||||| |||| |
Db      645 CACTTTGGGAGGCCAGGGTGGGCAGATCACCTGAGGTCGGGAGTTCGAGACCAGCCTGAC 704

Qy         157 CAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACA 216
               ||| ||| ||||||||||||| ||  |||||||||||||| |||| |||| ||||| ||
Db      705 CAATATGATGAAACCCCGTCTCTA-AAAAAATACAAAAATTAGCCGGGCGTGGTGTCAGG 763
```

```
Qy        217 CACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGA 276
              |||||||| |||||||||||||||||  | ||| ||||||||||||||| |||
Db        764 CACCTGTAATCCCAGCTACTCAGGAAGCAGAGGTTGCAGTGAGCTGAGGTCG-------- 815

Qy        277 GCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAA 336
              | |:|   |||      | | |  |  ||||||| ||:| |||||||||| || || |||
Db        816 GGCCATTGCACTCCAGCCTGGGCAAAAAGAGTGAAACTCTGTCTCAAAAAAAAAAAAAAA 875

Qy        337 AAAC 340
              ||||
Db        876 AAAC 879


RESULT 14
AAA64141
ID   AAA64141 standard; DNA; 14784 BP.
XX
AC   AAA64141;
XX
DT   20-DEC-2000  (first entry)
XX
DE   Nucleotide sequence of a beta-tubulin antigen.
XX
KW   Beta-tubulin antigen; inner ear protein; Meniere's disease; autoantibody;
KW   chronic ear disease; autoimmune disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200050593-A1.
XX
PD   31-AUG-2000.
XX
PF   25-FEB-2000; 2000WO-US004795.
XX
PR   25-FEB-1999;   99US-0121549P.
XX
PA   (UYTE-) UNIV TENNESSEE RES CORP.
XX
PI   Yoo TJ;
XX
DR   WPI; 2000-558400/51.
XX
PT   New beta-tubulin antigen in the membranous structure of the inner ear,
PT   reactive with antibodies of patients with Meniere's disease, for
PT   diagnosing Meniere's disease and distinguishing this disease from other
PT   autoimmune ear diseases.
XX
PS   Claim 3; Page 97-103; 115pp; English.
XX
CC   The present sequence encodes a beta-tubulin antigen. The protein is an
CC   antigen of the membranous structure of the inner ear protein, and is
CC   reactive with antibodies from patients having Meniere's disease.
CC   Meniere's disease is a chronic ear disease with unknown etiology. Serum
CC   from patients suffering from this disease contain autoantibodies against
CC   a 30 kDa cochlear protein antigen. The disease is believed to be an
CC   autoimmune disease. The beta-tubulin antigen is useful as a target
CC   substance in diagnosing or detecting Meniere's disease and in
CC   distinguishing this disease from other autoimmune ear diseases
XX
SQ   Sequence 14784 BP; 5454 A; 2966 C; 2926 G; 3438 T; 0 U; 0 Other;

  Query Match             38.3%;  Score 153.2;  DB 3;  Length 14784;
  Best Local Similarity   65.1%;  Pred. No. 2.9e-32;
  Matches  241;  Conservative    0;  Mismatches  128;  Indels    1;  Gaps    1;

Qy         31 GAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAA 90
              || |  | |  ||| ||  | | |||       |     |       ||   | ||||||||
Db       3372 GATCTAAGAATTTATTATAAAAGTGTAGCAAGGGCTGGGCGCGGTGCCTCACGCCTGTAA 3431

Qy         91 TCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCAT 150
              |||||||||| |||||||| |||||| |||||||||||||| ||| |||||||||
Db       3432 TCCCAGCACTTTGGGAGGCCAACGTGGGCAGATCACCTGAGGTCAGGAGTTCGAGACCAG 3491

Qy        151 CCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGT 210
              |||||||||||||| |||||||||| |||||||||||||||| ||||  ||||||||
Db       3492 CCTGGCCAACATGGCGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCAGGCATGGT 3551

Qy        211 GGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCA 270
              ||| |||||||||||||||||||||||||||||| |  ||   | ||        || | |
Db       3552 GGCGCACACCTGTAGTCCCAGCTACTCAGGAGGCTTAGGCAGGAGAACCACTTGAACCCG 3611

Qy        271 GAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAAC 330
              |  | |  |    ||  || |  ||| | ||  | |||     | | ||  | |||||
Db       3612 GGAGGTGTAGGTTGCAGTGAGCCGAGATCGTGC-CATTGCACGCCAGCCTGGACAACAAG 3670

Qy        331 AACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTATTGCAGGA 390
              | ||||||| |||||||     | || |   |       ||   | ||  |||||
Db       3671 AGCAAAAAAAAAAAAAAAATTGTGGCCAGTACAAAAACTTCATTATAATACATTTGCATTT 3730

Qy        391 GACCAAAAAG 400
                 |||  ||
Db       3731 TTCCATCCAG 3740


RESULT 15
AEE03362
ID   AEE03362 standard; DNA; 166942 BP.
XX
AC   AEE03362;
XX
DT   09-FEB-2006  (first entry)
XX
DE   Acute myelogenous leukemia prognosis related DNA sequence SEQ ID NO: 23.
XX
KW   ds; gene; acute myelogenous leukemia; prognosis; gene expression;
KW   biochip.
XX
OS   Homo sapiens.
XX
PN   JP2005333987-A.
XX
PD   08-DEC-2005.
XX
PF   06-MAY-2005; 2005JP-00135284.
XX
PR   06-MAY-2004; 2004US-0568635P.
XX
PA   (VERI-) VERIDEX LLC.
XX
PI   Laponi M;
XX
DR   WPI; 2006-004067/01.
XX
PT   Evaluating an acute myelogenous leukemia patient's prognosis, comprising
PT   detecting lower or higher expression level of gene recognized by probe
PT   set e.g. 202820-at and 206148-at, with respect to predetermined cut-off
PT   level.
XX
PS   Example 5; SEQ ID NO 23; 60pp; Japanese.
XX
CC   The present sequence is that of a human gene sequence which is claimed
CC   for use in evaluating the prognosis of patents suffering from acute
CC   myelogenous leukemia (AML) by analysis of the patients gene expression
CC   profile. The invention relates to a novel method for evaluating the
CC   prognosis of a patient with acute myelogenous leukemia by detecting
CC   higher/lower expression level of genes encoding mRNA recognized by a
CC   probe set chosen from 19 nucleotide sequences (AEE03340 or AEE03344-
CC   AEE03361). Also, gene expression profiling of AML patients using a set of
CC   probes based on 167 sequences (AEE03362-AEE03527) using a biochip was
CC   performed to determine prognosis. Methods are also included for
CC   determining an AML patient's state, determining AML patient's treatment
CC   protocol, determining whether the patient will respond to the treatment,
CC   and producing an AML patient's prognosis report by analysis of the
CC   patients gene expression profile. The methods are useful in evaluating an
CC   AML patient's prognosis and for treating AML patient.
XX
```

```
SQ   Sequence 166942 BP; 42646 A; 40900 C; 40852 G; 42544 T; 0 U; 0 Other;

  Query Match             38.2%;  Score 152.8;  DB 15;  Length 166942;
  Best Local Similarity   75.0%;  Pred. No. 7.8e-32;
  Matches  219;  Conservative    1;  Mismatches   63;  Indels    9;  Gaps    2;

Qy          84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
               ||| |||||||||||||| |||||||| |||  ||  |||||| ||||| || |||  ||
Db       33655 CCTATAATCCCAGCACTTTGGGAGGCCGAGGCAGGTAGATCACTTGAGGCCAGGAGTCCG 33714

Qy         144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
               |||||| |||||||||||||||||| ||| ||| ||||||||| ||||||| |||||||
Db       33715 AGACCAGCCTGGCCAACATGGTGAAATCCCATCTCTACTAAAAACACAAAAATTAGCTGG 33774

Qy         204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
               |||||||||  ||| |||||| |||||||||||  |||| | ||||||||||||||||||
Db       33775 GCATGGTGGTGCACGCCTGTAATCCCAGCTACTTGGGAGGCTGAGATTGCAGTGAGCTGA 33834

Qy         264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
               |||||           | |   ||| |   || | |  | |||||| ||||:|| ||||||
Db       33835 GATCG--------TGCCACTGCACTGCAGACTG-GGCAACAGAGTAAGACTCCATCTCAA 33885

Qy         324 AAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCC 375
               ||| || || |||| |   ||| |    || | | | |   ||| ||||| |
Db       33886 AAAAAATAAAAAAACAAGCAAACAAACAAACAAACTTTAGTTCTACGGTTGC 33937


RESULT 16
AAL04959/c
ID   AAL04959 standard; DNA; 4388 BP.
XX
AC   AAL04959;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Human reproductive system related antigen DNA SEQ ID NO: 7647.
XX
KW   Human; reproductive system related antigen; reproductive system disorder;
KW   cancer; gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155320-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001339.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241221P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
```

```
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465570/50.
XX
PT   Isolated nucleic acid molecule encoding a reproductive system antigen is
PT   used in preventing, treating or ameliorating a medical condition.
XX
PS   Disclosure; SEQ ID NO 7647; 1297pp + Sequence Listing; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human reproductive system related antigens. These can be used
CC   in the prevention and treatment of reproductive system disorders,
CC   including cancer. The present sequence is a genomic sequence encoding a
CC   protein of the invention
XX
SQ   Sequence 4388 BP; 870 A; 1238 C; 1205 G; 1075 T; 0 U; 0 Other;

  Query Match             38.1%;  Score 152.4;  DB 4;  Length 4388;
  Best Local Similarity   76.5%;  Pred. No. 3.3e-32;
  Matches  202;  Conservative    1;  Mismatches   52;  Indels    9;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | ||||||||| |||||||||  |||||||| |||  ||||||||||  ||||||||  |||
Db        1275 AAGCCTGTAATTCCAGCACTTTGGGAGGCCGAGGCCGGCGGATCACTTGAGGTCAGGAGT 1216

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               ||||||||| |||||||||||||||||||||| |||||| |||||||||||||||| ||||
Db        1215 TCGAGACCAGCCTGGCCAACATGGTGAAACTCCGTCTCTACTAAAAATACAAAAATTAGC 1156

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
                 | |||||||  | | ||||||  |||||||||||| |||| |||||  ||||||  ||||
Db        1155 CGCACGTGGTGGCGCGCGCCTGTAATCCCAGCTACTCGGGAGGCGGAGGTTGCAGCGAGC 1096

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               |||||||      |    |    |    |    |          || || |||||:||||||
Db        1095 AGAGATCGTGCCATTGCACTCCAGTCTGGGCGA---------CACAGCGAGACTCCGTCT 1045

Qy         321 CAAAAACAACAACAAAAAACAAAA 344
               |||||| || || |||||| ||||
Db        1044 CAAAAAAAAAAAAAAAAAAAAAA 1021


RESULT 17
ABL97853/c
ID   ABL97853 standard; DNA; 4388 BP.
XX
AC   ABL97853;
XX
DT   21-JUN-2002  (first entry)
XX
DE   Human testicular antigen encoding DNA fragment SEQ ID NO: 2505.
XX
KW   Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW   reproductive system disorder; urinary system disorder; gene therapy;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disease; infection; cytostatic; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155317-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001329.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
```

PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483232/52.
XX
PT Nucleic acids encoding 973 human testicular antigen polypeptides, useful
PT for preventing, diagnosing and/or treating testicular cancer.
XX
PS Disclosure; SEQ ID NO 2505; 766pp; English.
XX
CC The present invention provides the protein and coding sequences of 973
CC human testicular antigens, and fragments of their genomic sequences. The
CC sequences can be used in the treatment of cardiovascular, urinary system,
CC reproductive system, immune, respiratory, neurological and
CC gastrointestinal disorders, infections, and particularly cancer,
CC especially testicular cancers. The present sequence is a DNA encoding a
CC protein fragment of the invention
XX
SQ Sequence 4388 BP; 870 A; 1238 C; 1205 G; 1075 T; 0 U; 0 Other;

Query Match 38.1%; Score 152.4; DB 4; Length 4388;
Best Local Similarity 76.5%; Pred. No. 3.3e-32;
Matches 202; Conservative 1; Mismatches 52; Indels 9; Gaps 1;

```
Qy        81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
             | ||||||||| ||||||||| |||||||| |||  |||||||||| |||||||| |||
Db      1275 AAGCCTGTAATTCCAGCACTTTGGGAGGCCGAGGCCGGCGGATCACTTGAGGTCAGGAGT 1216

Qy       141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
             |||||||| ||||||||||||||||||| |||||| |||||||||||||||| ||||
Db      1215 TCGAGACCAGCCTGGCCAACATGGTGAAACTCCGTCTCTACTAAAAATACAAAAATTAGC 1156

Qy       201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
             |  | ||||||| | | |||||| |||||||||||| |||| ||||| |||||| ||||
Db      1155 CGCACGTGGTGGCGCGCGCCTGTAATCCCAGCTACTCGGGAGGCGGAGGTTGCAGCGAGC 1096

Qy       261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
             |||||||      |    |  |      |  |          || || |||||:||||||
Db      1095 AGAGATCGTGCCATTGCACTCCAGTCTGGGCGA---------CACAGCGAGACTCCGTCT 1045

Qy       321 CAAAAACAACAACAAAAAACAAAA 344
             |||||| || || |||||| ||||
Db      1044 CAAAAAAAAAAAAAAAAAAAAAAA 1021
```

RESULT 18
ABL68122/c
ID ABL68122 standard; DNA; 174424 BP.

XX
AC ABL68122;
XX
DT 15-MAY-2002 (first entry)
XX
DE Ovary cancer related gene sequence SEQ ID NO:6459.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 6459; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 174424 BP; 39582 A; 48304 C; 48535 G; 38003 T; 0 U; 0 Other;

Query Match 38.0%; Score 152; DB 6; Length 174424;
Best Local Similarity 71.8%; Pred. No. 1.3e-31;
Matches 214; Conservative 0; Mismatches 80; Indels 4; Gaps 1;

```
Qy          6 TACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGG 65
              | |||||   | || ||  || |||     |     ||   |||||    | |||||  ||
Db      48438 TCCTCAGGTGTCTGGTGATCCTTGGTTGTCTTCTCATATTTAAGAATGGGGCAGGCTGGG 48379

Qy         66 CATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCA 125
              |||| ||        |||||||||||||| ||||| |||||||| | ||||||||||||
Db      48378 CATGGCA----GCTCATGCCTGTAATCCCAACACTTTGGGAGGCCGAAGTGGGCGGATCA 48323

Qy        126 CCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAA 185
              |||||||||||||||| || |||||||||||||||||||||||| || | |||||||
Db      48322 CCTGAGGTCAAGAGATCAAGTCCATCCTGGCCAACATGGTGAAACCCTGTTTGTACTAAA 48263

Qy        186 AATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCG 245
              |||||||||| |||||||||||||| || |||||||||||||||||||  ||||||
Db      48262 AATACAAAAATTAGCTGGGCATGGTGGTACGCACCTGTAGTCCCAGCTACTTGGGAGCCT 48203

Qy        246 GAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
              |||   | ||         ||   | |   | |  ||   ||  || |  ||| || ||
Db      48202 GAGGCAGGAGAATCGCTTGAATCCGGGAGGTGGAGATTGCAGTGAGCCGAGATTGTGC 48145
```

RESULT 19
ADQ19573/c
ID ADQ19573 standard; DNA; 181343 BP.
XX
AC ADQ19573;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 2392.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX

```
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 2392; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   DNA of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 181343 BP; 41465 A; 49982 C; 50309 G; 39587 T; 0 U; 0 Other;

  Query Match          38.0%;  Score 152;  DB 12;  Length 181343;
  Best Local Similarity  71.8%;  Pred. No. 1.3e-31;
  Matches  214;  Conservative    0;  Mismatches   80;  Indels    4;  Gaps    1;

Qy          6 TACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGG 65
              | |||||   | ||  ||  || |||    |     ||   |||||    | |||||  ||
Db      48439 TCCTCAGGTGTCTGGTGATCCTTGGTTGTCTTCTCATATTTAAGAATGGGGCAGGCTGGG 48380

Qy         66 CATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCA 125
              ||||  ||         |||||||||||||||| ||||| |||||||| | ||||||||||||
Db      48379 CATGGCA----GCTCATGCCTGTAATCCCAACACTTTGGGAGGCCGAAGTGGGCGGATCA 48324

Qy        126 CCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAA 185
              |||||||||||||||| || |||||||||||||||||||||||||| || | |||||||
Db      48323 CCTGAGGTCAAGAGATCAAGTCCATCCTGGCCAACATGGTGAAACCCTGTTTGTACTAAA 48264

Qy        186 AATACAAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCG 245
              |||||||||| |||||||||||||||| || |||||||||||||||||||||  ||||||
Db      48263 AATACAAAAATTAGCTGGGCATGGTGGTACGCACCTGTAGTCCCAGCTACTTGGGAGCCT 48204

Qy        246 GAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
              |||   | ||         ||   | |   | |  ||   ||  || |  ||| || ||
Db      48203 GAGGCAGGAGAATCGCTTGAATCCGGGAGGTGGAGATTGCAGTGAGCCGAGATTGTGC 48146


RESULT 20
ABQ88186/c
ID   ABQ88186 standard; cDNA; 169739 BP.
XX
AC   ABQ88186;
XX
DT   18-SEP-2002  (first entry)
XX
DE   Human osteoblast differentiation related cDNA SEQ ID NO 93.
XX
KW   Human; osteoblast; stem cell differentiation; bone tissue deposition;
KW   osteoporosis; osteopathic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200250301-A2.
XX
PD   27-JUN-2002.
XX
PF   18-DEC-2001; 2001WO-US048276.
XX
PR   18-DEC-2000; 2000US-0255882P.
PR   24-APR-2001; 2001US-0285691P.
XX
PA   (GENE-) GENE LOGIC INC.
PA   (PROC ) PROCTER & GAMBLE CO.
XX
PI   Ji D,  Axelrod DW,  Cook JS,  Jaiswal N,  Einstein R,  Houghton A;
PI   Mertz L;
XX
DR   WPI; 2002-557663/59.
XX
PT   Use of genes and their expression profiles associated with osteoblast
PT   differentiation for screening modulators bone formation, for diagnosing
PT   or treating e.g. osteoporosis, or as markers for the differentiation
PT   process.
XX
PS   Claim 1; SEQ ID NO 93; 78pp + Sequence Listing; English.
XX
CC   The invention relates to genes and their expression profiles are used
CC   for: (a) screening modulators of precursor stem cell differentiation into
CC   osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
CC   deposition of bone tissue, abnormal rate of osteoblast formation or
CC   osteoporosis; or (c) treating or monitoring treatment of the conditions
CC   cited in (b), or monitoring the progression of bone tissue deposition.
CC   Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC   osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
CC   induced abnormalities in bone formation or bone loss, conditions that
CC   involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
CC   skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
CC   or fibrous dysplasia. The present sequence is that of an osteoblast
CC   differentiation associated cDNA marker of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 169739 BP; 49809 A; 35660 C; 35715 G; 48555 T; 0 U; 0 Other;

  Query Match          37.9%;  Score 151.4;  DB 6;  Length 169739;
  Best Local Similarity  70.9%;  Pred. No. 1.9e-31;
  Matches  229;  Conservative    1;  Mismatches   87;  Indels    6;  Gaps    2;

Qy         30 GGAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTA 89
              || |   |||| ||| | | | ||    ||    | || |      |   || |||||||
Db     125892 GGTAAGAAAATCTTAGTTAAAGATCTAAAGTTTGGCCAGGCACAGTGGCTAACGCCTGTA 125833

Qy         90 ATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCA 149
              ||||||||||||||||||||||||||  || ||||||||||||||| ||| |||||||||
Db     125832 ATCCCAGCACTTCGGGAGGCCAAGGCAGGTGGATCACCTGAGGTCAGGAGTTCGAGACCA 125773

Qy        150 TCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGG 209
              |||||||||||||||||||||| |||| |||||||||||||||| |||| ||||||||
Db     125772 GCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCCGGGCATGG 125713

Qy        210 TGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGC 269
              ||||| |||||||||||||| ||||||| ||| | |||   | ||         || | |
Db     125712 TGGCAGACACCTGTAGTCCCAACTACTCAAGAGGCTGAGGCAGGAGAATCACTTGAAC-C 125654

Qy        270 AGAGTGAGCCGAAATCACA-----GATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAA 324
              | ||||| | | ||| ||      |  | |       |   |  ||||||:|| |||| ||
Db     125653 TGGGTGAGTCAAGATCGCAACATGGCACTCCAGCCTGGGCAACAGAGACTCCATCTCCAA 125594

Qy        325 AACAACAACAAAAAACAAAAAAA 347
              || || ||  ||| | |||| ||
Db     125593 AAAAATAAATAAATAAAAAATAA 125571


RESULT 21
AAK68418/c
ID   AAK68418 standard; DNA; 14282 BP.
XX
AC   AAK68418;
XX
DT   06-NOV-2001  (first entry)
```

XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23230.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.

```
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS   Disclosure; SEQ ID NO 23230; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting the
CC   nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention
XX
SQ   Sequence 14282 BP; 3993 A; 2449 C; 2814 G; 5026 T; 0 U; 0 Other;

  Query Match             37.7%;  Score 150.6;  DB 4;  Length 14282;
  Best Local Similarity   73.6%;  Pred. No. 1.5e-31;
  Matches  192;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          48 AGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAG 107
               | | | ||| |  |  | || |     ||    ||||||||||||||||||||| |||||
Db        4784 AAATAATGTTATCCTGGCCAGGCACGGGGGCACATGCCTGTAATCCCAGCACTTTGGGAG 4725

Qy         108 GCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGA 167
                || |||||||| |||||| |||||||| ||| | ||| ||| ||||||| ||||||| |
Db        4724 ACCGAGGTGGGCAGATCACTTGAGGTCAGGAGTTTGAGGCCAGCCTGGCCGACATGGTAA 4665

Qy         168 AACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTC 227
               |||||||||| |||||||||| ||||| ||||||||   | ||||||  ||||| |||||
Db        4664 AACCCCGTCTCTACTAAAAATAAAAAAATTAGCTGGGTGTCGTGGCATGCACCTATAGTC 4605

Qy         228 CCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCAC 287
               |||||||| ||||| | ||| |||||||||| |||||| ||           | |
Db        4604 CCAGCTACTGAGGAGGCAGAGGTTGCAGTGAGCCGAGATCACACCACTGCATTCCAGCCT 4545

Qy         288 AGATCACAGAGTGAGCAGAGT 308
                |   ||||||||| |   ||
Db        4544 GGGCAACAGAGTGAACTCTGT 4524


RESULT 22
AAS28363/c
ID   AAS28363 standard; DNA; 32146 BP.
XX
AC   AAS28363;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Genomic sequence #203 encoding for novel human respiratory antigen.
XX
KW   Human; respiratory antigen; respiratory disorder; throat disorder;
KW   lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
KW   anti allergic; anti asthmatic; anti inflammatory; olfactory;
KW   respiratory active; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155448-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001333.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
```

PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-476224/51.
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the respiratory system including respiratory cancers
PT and also for testing and detection e.g. diagnosis.
XX
PS Disclosure; SED ID No 797; 546pp; English.
XX
CC The present invention relates to the isolation of novel human respiratory
CC antigens (AAU17685-AAU17975), and cDNA and genomic sequences encoding for
CC these polypeptides. The sequences of the invention are useful for
CC preventing, treating and/or prognosing disorders related to the
CC respiratory system including throat disorders (e.g. vocal cord paralysis,
CC tonsillitis, and laryngitis), lung disorders e.g. pneumonia, allergic
CC disorders e.g. asthma, pleurisy, cystic fibrosis, emphysema, nose
CC disorders and cancers of the respiratory tissues e.g. lung cancer. The
CC polynucleotide sequences of the invention are useful in gene therapy and
CC antisense therapy. AAS28161-AAS28764 represent genomic sequences encoding
CC for novel human respiratory antigens. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 32146 BP; 8745 A; 6747 C; 7148 G; 9506 T; 0 U; 0 Other;

Query Match 37.7%; Score 150.6; DB 4; Length 32146;
Best Local Similarity 68.5%; Pred. No. 1.9e-31;
Matches 224; Conservative 1; Mismatches 95; Indels 7; Gaps 1;

```
Qy        81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
             | ||||||||| | ||||||| |||||||||||| ||| ||||||| |||||||| |||
Db      2437 ACGCCTGTAATGCTAGCACTTTGGGAGGCCAAGGCGGGAGGATCACTTGAGGTCAGGAGT 2378

Qy       141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
             | ||||| | |||||||||||||||||||||| |||| |||||||||||||||| ||||
Db      2377 TTGAGACAAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGC 2318

Qy       201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAG-------CCGGAGATTGC 253
              |||||||||||   |||||||||||||||| ||  ||||       | ||||| |
Db      2317 AGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTGCTTGGGAGACTGAGCCAGGAGAATTG 2258

Qy       254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGAC 313
               ||| |   |   || |||   ||    | |  ||||   |        |    |||||
Db      2257 CTTGAACCCGGGAGGCGGAGGTTGCAATGAGCCGAGATTGTACTACTGCAATGCAGAGAC 2198

Qy       314 KCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTT 373
             :|||||||||||| || || |||||| |||||||  |||| |       ||   | ||
Db      2197 TCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGATAAAAGTAGAGACAATAGGGGGA 2138

Qy       374 CCCAGACTATTGCAGGAGACCAAAAAG 400
               |   | | | | |||    | ||||
Db      2137 TCTCAGCAAATACTGGATTTAACAAAG 2111
```

RESULT 23
ADG41559/c
ID ADG41559 standard; DNA; 32146 BP.
XX
AC ADG41559;
XX
DT 26-FEB-2004 (first entry)
XX
DE Human respiratory system associated genomic DNA seq id 797.
XX
KW antiinflammatory; antiallergic; antiasthmatic; cytostatic; gene therapy;
KW respiratory system antigen;
KW human respiratory system associated polynucleotide;

KW respiratory system disorder; throat disorder; vocal cord paralysis;
KW tonsillitis; laryngitis; lung disorder; pneumonia; allergic disorder;
KW asthma; eosinophilic pneumonia; pleurisy; cystic fibrosis; emphysema;
KW histiocytosis; sarcoidosis; nose disorder; rhinitis; sinusitis; neoplasm;
KW cancer; respiratory tissue cancer; throat cancer; lung cancer;
KW cancer of the nose; gene therapy; chromosome identification; forensic;
KW human respiratory system associated protein; ds; human.
XX
OS Homo sapiens.
XX
PN US2003215893-A1.
XX
PD 20-NOV-2003.
XX
PF 07-AUG-2002; 2002US-00212872.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 28-SEP-2000; 2000US-0235935P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.

```
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
PR   17-JAN-2001; 2001US-00764860.
PR   14-FEB-2002; 2002US-00074095.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM,  Barash SC;
XX
DR   WPI; 2003-902033/82.
XX
PT   Novel respiratory system antigen and polynucleotides encoding the
PT   polypeptides; useful for treating diagnosing, treating or preventing
PT   tonsillitis, pneumonia, asthma and cystic fibrosis, emphysema, throat
PT   cancer.
XX
PS   Disclosure; SEQ ID NO 797; 236pp; English.
XX
CC   The invention describes an isolated polypeptide (I) comprising an amino
CC   acid sequence that is at least 90% identical to polypeptide fragment of
CC   any one of 299 respiratory system antigen sequences (PS) and having
CC   biological activity, polypeptide domain or epitope of PS, full-length
CC   protein of PS, or variant, allelic variant or species homolog of PS. (I)
CC   or a polynucleotide (II) encoding (I) is also useful for diagnosing a
CC   pathological condition or a susceptibility to a pathological condition in
CC   a subject which involves determining the presence or absence of mutation
CC   in (II) or determining the presence or amount of expression of (I) in a
CC   biological sample and diagnosing a pathological condition based on the
CC   result. The human respiratory system associated polynucleotides, the
CC   polypeptides encoded by them, and antibodies that immunospecifically bind
CC   these polypeptides are useful in diagnosis, treatment, prevention and/or
CC   prognosis of disorders of respiratory system such as throat disorders
CC   (e.g., vocal cord paralysis, tonsillitis, and laryngitis), lung disorders
CC   (e.g., pneumonia), allergic disorders, (e.g., asthma and eosinophilic
CC   pneumonia), pleurisy, cystic fibrosis, emphysema, histiocytosis,
CC   sarcoidosis, nose disorders (rhinitis and sinusitis), neoplasms and/or
CC   cancers of respiratory tissues (e.g., throat cancer, lung cancer, and
CC   cancer of the nose). The polynucleotides are useful in gene therapy
CC   techniques, for chromosome identification, identifying individuals from

  Query Match             37.7%;  Score 150.6;  DB 10;  Length 32146;
  Best Local Similarity   68.5%;  Pred. No. 1.9e-31;
  Matches  224;  Conservative    1;  Mismatches   95;  Indels    7;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||| | ||||||| |||||||||||| ||| ||||||| |||||||| |||
Db       2437 ACGCCTGTAATGCTAGCACTTTGGGAGGCCAAGGCGGGAGGATCACTTGAGGTCAGGAGT 2378

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||| | |||||||||||||||||||||| |||| ||||||||||||||||| ||||
Db       2377 TTGAGACAAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGC 2318

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAG-------CCGGAGATTGC 253
              ||||||||||||   |||||||||||||||||| ||  ||||       | |||||  |
Db       2317 AGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTGCTTGGGAGACTGAGCCAGGAGAATTG 2258

Qy        254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGAC 313
              ||| |   |   || |||   ||    | |  ||||   |         |    |||||
Db       2257 CTTGAACCCGGGAGGCGGAGGTTGCAATGAGCCGAGATTGTACTACTGCAATGCAGAGAC 2198

Qy        314 KCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTT 373
              :||||||||||| || || |||||| |||||||  |||| |         ||   | ||
Db       2197 TCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGATAAAAGTAGAGACAATAGGGGGA 2138

Qy        374 CCCAGACTATTGCAGGAGACCAAAAAG 400
              |    | | | | |||    | ||||
Db       2137 TCTCAGCAAATACTGGATTTAACAAAG 2111
```

```
RESULT 24
ADI97333/c
ID   ADI97333 standard; DNA; 32146 BP.
XX
AC   ADI97333;
XX
DT   04-NOV-2004  (first entry)
XX
DE   Human respiratory system associated polypeptide-related DNA SeqID797.
XX
KW   respiratory system-related polypeptide; antiasthmatic; antibacterial;
KW   antiinflammatory; cytostatic; antianaemic; antiallergic; gene therapy;
KW   pneumonia; lung cancer; cystic fibrosis; asthma; sarcoidosis; rhinitis;
KW   anaemia; leukaemia; inflammation; sinusitis;
KW   chronic obstructive pulmonary disease; infectious disease; human; ds.
XX
OS   Homo sapiens.
XX
PN   US2003077704-A1.
XX
PD   24-APR-2003.
XX
PF   14-FEB-2002; 2002US-00074095.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
```

PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764860.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2003-765403/72.
XX
PT New human respiratory system-related polypeptide and genes, useful for
PT treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic
PT fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or
PT sinusitis.
XX
PS Disclosure; SEQ ID NO 797; 202pp; English.
XX
CC This invention is related to a novel isolated polypeptide, which
CC comprises a human respiratory system-related polypeptide, and the DNA
CC sequence which encodes it. The invention may be useful for the
CC development of compounds with an antiasthmatic, antibacterial,
CC antiinflammatory, cytostatic, antianaemic or antiallergic activity. In
CC addition, the sequences disclosed may be useful for gene therapy. The
CC polypeptide or polynucleotide is useful for treating, preventing or
CC ameliorating a medical condition, for example pneumonia, lung cancer,
CC cystic fibrosis, asthma, sarcoidosis, rhinitis, anaemia, leukaemia,
CC inflammations, sinusitis, chronic obstructive pulmonary disease or
CC infectious diseases. The polypeptide or polynucleotide is also useful for
CC diagnosing any of these diseases or a susceptibility to the disease. The
CC present sequence is that of a human DNA sequence which is related to a
CC human respiratory system associated gene of the invention.
XX
SQ Sequence 32146 BP; 8745 A; 6747 C; 7148 G; 9506 T; 0 U; 0 Other;

Query Match 37.7%; Score 150.6; DB 11; Length 32146;
Best Local Similarity 68.5%; Pred. No. 1.9e-31;
Matches 224; Conservative 1; Mismatches 95; Indels 7; Gaps 1;

```
Qy        81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
             | |||||||||| | ||||||| |||||||||||| ||| ||||||| |||||||| |||
Db      2437 ACGCCTGTAATGCTAGCACTTTGGGAGGCCAAGGCGGGAGGATCACTTGAGGTCAGGAGT 2378

Qy       141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
             | ||||| | |||||||||||||||||||||| |||| |||||||||||||||||| ||||
Db      2377 TTGAGACAAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGC 2318

Qy       201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAG-------CCGGAGATTGC 253
              ||||||||||||    |||||||||||||||| ||  ||||       | ||||| |
Db      2317 AGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTGCTTGGGAGACTGAGCCAGGAGAATTG 2258

Qy       254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGAC 313
               ||| |   |  || |||   ||    | |  ||||  |       |    |||||
Db      2257 CTTGAACCCGGGAGGCGGAGGTTGCAATGAGCCGAGATTGTACTACTGCAATGCAGAGAC 2198

Qy       314 KCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTT 373
             :|||||||||||| || || |||||| |||||||  |||| |       ||   | ||
Db      2197 TCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAGATAAAAGTAGAGACAATAGGGGGA 2138

Qy       374 CCCAGACTATTGCAGGAGACCAAAAAG 400
               |  | | | |||    | ||||
Db      2137 TCTCAGCAAATACTGGATTTAACAAAG 2111
```

```
RESULT 25
ADL13941/c
ID   ADL13941 standard; DNA; 125515 BP.
XX
AC   ADL13941;
XX
DT   06-MAY-2004 (first entry)
XX
DE   Osteoarthritis-associated polymorphic nucleotide #473.
XX
KW   ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW   joint space narrowing; osteophyte development; joint pain;
KW   osteoarthritis; SNP; single nucleotide polymorphism.
XX
OS   Homo sapiens.
XX
PN   WO2003054166-A2.
XX
PD   03-JUL-2003.
XX
PF   19-DEC-2002; 2002WO-US041225.
XX
PR   20-DEC-2001; 2001US-0342603P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Jones KA, Schafer A;
XX
DR   WPI; 2003-559141/52.
XX
PT   Determining susceptibility of an individual to joint space narrowing,
PT   osteophyte development and/or joint pain comprises identifying whether
PT   the individual has at least one polymorphism in a polynucleotide encoding
PT   a protein.
XX
PS   Disclosure; SEQ ID NO 473; 297pp; English.
XX
CC   The invention relates to a method of determining susceptibility of an
CC   individual to joint space narrowing and/or osteophyte development and/or
CC   joint pain comprising identifying whether the individual has at least one
CC   polymorphism in a polynucleotide encoding at least one of the protein
CC   listed in the specification. The methods, composition and agent are
CC   useful for modulating the susceptibility of an individual to joint space
CC   narrowing and/or osteophyte development and/or joint pain that is
CC   associated with a disease, preferably osteoarthritis. The cell line and
CC   the non-human animal are useful for screening for an agent for diagnosing
CC   an individual having susceptibility to joint space narrowing and/or
CC   osteophyte development and/or joint pain. This sequence corresponds to
CC   the polynucleotide encoding a protein listed in the specification. (Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences).
XX
SQ   Sequence 125515 BP; 33180 A; 28822 C; 28744 G; 34769 T; 0 U; 0 Other;

  Query Match           37.7%;  Score 150.6;  DB 10;  Length 125515;
  Best Local Similarity 75.7%;  Pred. No. 3e-31;
  Matches  202;  Conservative    1;  Mismatches   55;  Indels    9;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||| |||||||||||||||| |||||||  ||| ||||||||  |  |||||||| |||
Db      52801 ATGCTTGTAATCCCAGCACTTTGGGAGGCTGAGGAGGGCGGATTATTTGAGGTCAGGAGT 52742

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              |  |||||| |||||||||||||||||||||||| |||| ||||||||||||||||| ||||
Db      52741 TTCAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGC 52682

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||| ||||||||| |||||||||||||||| |||||||  |  || ||||||||||
Db      52681 CGGGCCTGGTGGCACGCACCTGTAGTCCCAGCTGCTCAGGAGACACAGGTTGCAGTGAGC 52622

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||||            |   | |   |   |          ||||| |||||:||  ||
Db      52621 CAAGATCATGCCACTGCACTCCAGCCTGGGCGA---------CAGAGCGAGACTCCACCT 52571

Qy        321 CAAAAACAACAACAAAAAACAAAAAAA 347
              |||||| || || |||||| ||||| |
Db      52570 CAAAAAAAAAAAAAAAAAAAAAAAAGA 52544


RESULT 26
AAZ29204/c
ID   AAZ29204 standard; DNA; 17538 BP.
XX
AC   AAZ29204;
XX
DT   21-FEB-2000  (first entry)
XX
DE   Human myelin oligodendrocyte glycoprotein gene.
XX
KW   Human; myelin oligodendrocyte glycoprotein; MOG; NS-specific antigen;
KW   nervous system-specific antigen; T cell; peripheral nervous system; PNS;
KW   central nervous system; CNS; nerve regeneration; neuronal degeneration;
KW   spinal cord injury; blunt trauma; penetrating trauma; senile dementia;
KW   ischaemic stroke; diabetic neuropathy; glaucoma; haemorrhagic stroke;
KW   Alzheimer's disease; Parkinson's disease; Huntington's chorea;
KW   amyotrophic lateral sclerosis; ALS; treatment; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1166. .15142
FT                   /*tag= a
FT                   /product= "myelin oligodendrocyte glycoprotein"
FT                   /note= "Reading frame is interrupted by introns the
FT                   precise location of which is not given in the
FT                   specification "
XX
PN   WO9960021-A2.
XX
PD   25-NOV-1999.
XX
PF   19-MAY-1999;   99WO-US010953.
XX
PR   19-MAY-1998;   98IL-00124550.
PR   21-JUL-1998;   98WO-US014715.
PR   22-DEC-1998;   98US-00218277.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
PA   (MCIN/) MCINNIS P A.
XX
PI   Eisenbach-Schwartz M,  Cohen IR,  Beserman P,  Mosonego A,  Moalem G;
XX
DR   WPI; 2000-072430/06.
DR   P-PSDB; AAY44236.
XX
PT   New compositions useful to treat nervous system injury or disease e.g.
PT   traumatic injury, Alzheimer's disease etc.
XX
PS   Claim 9; Fig 18; 92pp; English.
XX
CC   The present sequence is a gene encoding human myelin oligodendrocyte
CC   glycoprotein which is a nervous system-specific antigen. The antigen or
CC   peptides derived from it activate T cells in vivo. The present sequence
CC   is used to promote nerve regeneration or to prevent or inhibit neuronal
CC   degeneration caused by injury or diseases of nerves within the CNS or
CC   PNS. Such injury includes spinal cord injury, blunt trauma, penetrating
CC   trauma, haemorrhagic stroke or ischaemic stroke, whilst diseases include
CC   diabetic neuropathy, senile dementia, Alzheimer's disease, Parkinson's
CC   disease, glaucoma, Huntington's chorea, amyotrophic lateral sclerosis,
CC   etc
XX
SQ   Sequence 17538 BP; 4624 A; 4121 C; 3990 G; 4803 T; 0 U; 0 Other;
```

```
  Query Match             37.6%;  Score 150.4;  DB 3;  Length 17538;
  Best Local Similarity   72.1%;  Pred. No. 1.8e-31;
  Matches  196;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               ||||||||||||||||||||| || |||||||||  ||| ||||||||||||||| | ||
Db       13528 ATGCCTGTAATCCCAGCACTTTGGAAGGCCAAGGCAGGCAGATCACCTGAGGTCAGGGGT 13469

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               || |||||| |||||||||||||||||||||||||| |||||||||||||||| ||||
Db       13468 TCAAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 13409

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||||||||  | |||||||| ||||| ||| |||| | |||   | ||
Db       13408 TGGGCATGGTGGCATGCGCCTGTAGTTCCAGCAACTTGGGAGGCTGAGGCAGGAGAATCG 13349

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                 || | |||   | |  |    ||  || |  ||| || ||      |  |    ||
Db       13348 CTTGAACCCAGGAGGTGGAGGTTGCAGTGAGCCGAGATTGTGCCATTGCACTCCATCTCA 13289

Qy         321 CAAAAACAACAACAAAAAACAAAAAAACCATA 352
               ||||||  || |  |||| |||| ||  | |
Db       13288 AAAAAACTTCATCTCAAAAAAAAAGAAAGAAA 13257


RESULT 27
ABN86597/c
ID   ABN86597 standard; DNA; 17538 BP.
XX
AC   ABN86597;
XX
DT   05-NOV-2002  (first entry)
XX
DE   Human myelin oligodendrocyte glycoprotein (MOG) gene sequence.
XX
KW   Nerve regeneration; neuroprotection; neuronal degeneration; CNS; PNS;
KW   central nervous system; peripheral nervous system; tranquillizer; MOG;
KW   vulnerary; cerebroprotective; anti-tumour; antidiabetic; anticonvulsant;
KW   nootropic; antiparkinsonian; ophthalmological; analgesic; hepatotropic;
KW   osteopathic; vasotropic; nephrotropic; cytostatic; antigen; gene therapy;
KW   myelin oligodendrocyte glycoprotein; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   US2002072493-A1.
XX
PD   13-JUN-2002.
XX
PF   28-JUN-2001; 2001US-00893348.
XX
PR   19-MAY-1998;   98IL-00124500.
PR   21-JUL-1998;   98WO-US014715.
PR   22-DEC-1998;   98US-00218277.
PR   19-MAY-1999;   99US-00314161.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Eisenbach-Schwartz M,  Hauben E,  Cohen IR,  Beserman P,  Mosonego A;
PI   Moalem G;
XX
DR   WPI; 2002-607255/65.
DR   P-PSDB; ABB81071.
XX
PT   Promoting nerve regeneration and preventing neuronal degeneration in the
PT   central/peripheral nervous system from injury/disease, comprises
PT   administering nervous system-specific activated T cells/antigen, or
PT   analogs/peptides.
XX
PS   Disclosure; Fig 18; 93pp; English.
XX
CC   The invention relates to promoting nerve regeneration or conferring
CC   neuroprotection and preventing or inhibiting neuronal degeneration in the
CC   central/peripheral nervous system (NS). The method involves administering
CC   NS-specific activated T cells, NS-specific antigen, its analogue or its
CC   peptide, a nucleotide sequence the NS-specific antigen or its analogue or
CC   combinations. The method is useful for promoting nerve regeneration and
CC   preventing neuronal degeneration in central/peripheral nervous system
CC   from injury/disease, where the injury is spinal cord injury, blunt
CC   trauma, penetrating trauma, hemorrhagic stroke, ischaemic stroke or
CC   damages caused by surgery such as tumour excision. The disease is not an
CC   autoimmune disease or neoplasm. The disease results in a degenerative
CC   process occurring in either gray or white matter or both. The disease is
CC   diabetic neuropathy, senile dementia, Alzheimer's disease, Parkinson's
CC   disease, facial nerve (Bell's) palsy, glaucoma, Huntington's chorea,
CC   amyotrophic lateral sclerosis, non-arteritic optic neuropathy, and
CC   vitamin deficiency, intervertebral disc herniation, prion diseases such
CC   as Creutzfeldt-Jakob disease, carpal tunnel syndrome, peripheral
CC   neuropathies associated with various diseases, including but not limited
CC   to uremia, porphyria, hypoglycemia, Sjorgren Larsson syndrome, acute
CC   sensory neuropathy, chronic ataxic neuropathy, biliary cirrhosis, primary
CC   amyloidosis, obstructive lung diseases, acromegaly, malabsorption
CC   syndromes, polycythemia vera, immunoglobulin (Ig)A- and IgG gamma-
CC   pathies, complications of various drugs (e.g., metronidazole) and toxins
CC   (e.g., alcohol or organophosphates), Charcot-Marie-Tooth disease, ataxia
CC   telangectasia, Friedreich's ataxia, amyloid polyneuropathies,
CC   adrenomyeloneuropathy, Giant axonal neuropathy, Refsum's disease, Fabry's
CC   disease, or lipoproteinemia. The present sequence represents a human
CC   myelin oligodendrocyte glycoprotein (MOG) gene sequence, an example of NS
CC   -specific antigen
XX
SQ   Sequence 17538 BP; 4624 A; 4120 C; 3991 G; 4803 T; 0 U; 0 Other;

  Query Match             37.6%;  Score 150.4;  DB 6;  Length 17538;
  Best Local Similarity   72.1%;  Pred. No. 1.8e-31;
  Matches  196;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               ||||||||||||||||||||| || |||||||||  ||| ||||||||||||||| | ||
Db       13528 ATGCCTGTAATCCCAGCACTTTGGAAGGCCAAGGCAGGCAGATCACCTGAGGTCAGGGGT 13469

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               || |||||| |||||||||||||||||||||||||| |||||||||||||||| ||||
Db       13468 TCAAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 13409

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||||||||  | |||||||| ||||| ||| |||| | |||   | ||
Db       13408 TGGGCATGGTGGCATGCGCCTGTAGTTCCAGCAACTTGGGAGGCTGAGGCAGGAGAATCG 13349

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                 || | |||   | |  |    ||  || |  ||| || ||      |  |    ||
Db       13348 CTTGAACCCAGGAGGTGGAGGTTGCAGTGAGCCGAGATTGTGCCATTGCACTCCATCTCA 13289

Qy         321 CAAAAACAACAACAAAAAACAAAAAAACCATA 352
               ||||||  || |  |||| |||| ||  | |
Db       13288 AAAAAACTTCATCTCAAAAAAAAAGAAAGAAA 13257


RESULT 28
ADQ97695/c
ID   ADQ97695 standard; DNA; 88892 BP.
XX
AC   ADQ97695;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Human cancer associated sequence HD10-029, SEQ ID 672.
XX
KW   Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX
OS   Homo sapiens.
XX
PN   WO2004060304-A2.
XX
PD   22-JUL-2004.
XX
PF   22-DEC-2003; 2003WO-US041389.
```

XX
PR 27-DEC-2002; 2002US-00330773.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Malandro MS;
XX
DR WPI; 2004-543781/52.
XX
PT New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.
XX
PS Claim 1; SEQ ID NO 672; 199pp; English.
XX
CC The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic formate directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 88892 BP; 25748 A; 16788 C; 17057 G; 27074 T; 0 U; 2225 Other;

Query Match 37.6%; Score 150.4; DB 12; Length 88892;
Best Local Similarity 74.5%; Pred. No. 3e-31;
Matches 204; Conservative 1; Mismatches 62; Indels 7; Gaps 1;

```
Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| ||||||| ||| ||| ||||||||||||||| |||
Db      58516 ACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGTGGATCACCTGAGGTCAGGAGT 58457

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              ||||||||| |||| |||||||||||||||||| ||| ||||||||||||||||| ||||
Db      58456 TCGAGACCAGCCTGACCAACATGGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGCC 58397

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-------GCCGGAGATTGC 253
              |||||||||||||||| |||||| | |||||||||||||       || ||||| |
Db      58396 TGGGCATGGTGGCACACGCCTGTAATTCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTG 58337

Qy        254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGAC 313
              |||           | |  | |   |||       |  || ||  ||||||||||
Db      58336 CTTGAATAAGTGAGCTGAGATCATGCCACTGCACTCCAGGCTCAGCGACAAGAGTGAGAC 58277

Qy        314 KCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
              :  |||||||||| || || | |||| |||||||
Db      58276 TTTGTCTCAAAAAAAAAAAAAGAAAAGAAAAAAA 58243
```

RESULT 29
AAI61042
ID AAI61042 standard; cDNA; 4779 BP.
XX
AC AAI61042;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 5031.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM41886.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 5031; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 4779 BP; 1453 A; 960 C; 1048 G; 1318 T; 0 U; 0 Other;

Query Match 37.5%; Score 150; DB 4; Length 4779;
Best Local Similarity 79.7%; Pred. No. 1.6e-31;
Matches 177; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

```
Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| ||||||||||||  ||| |||||| ||||||||| |||
Db       4253 ACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCTGGCAGATCACTTGAGGTCAGGAGT 4312

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              || |||||| |||||||||||||||||||||||| ||| ||||||||||||||||| ||||
Db       4313 TCAAGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 4372

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||| ||||||||||  |||||||||||||||| |  |||| | |||  ||||||||||
Db       4373 TGGGCGTGGTGGCACATGCCTGTAGTCCCAGCTATTTGGGAGGCTGAGGCTGCAGTGAGC 4432

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAG 302
              |||||  ||          |   | |   |   ||| ||||||
Db       4433 CGAGATTGCGCCACTGCACTCCAGCCTGGGCGACACAGTGAG 4474
```

RESULT 30
AED18472
ID AED18472 standard; DNA; 100998 BP.
XX
AC AED18472;
XX
DT 15-DEC-2005 (first entry)
XX
DE Fibrotic disorder associated polynucleotide SEQ ID NO 723.

XX
KW antiinflammatory; gene therapy; fibrogenesis; gene expression;
KW therapeutic; diagnosis; uterine fibroids; gynecological; inflammation;
KW ds.
XX
OS Homo sapiens.
XX
PN WO2005098041-A2.
XX
PD 20-OCT-2005.
XX
PF 28-MAR-2005; 2005WO-US010257.
XX
PR 26-MAR-2004; 2004US-0556546P.
PR 19-OCT-2004; 2004US-0620444P.
PR 15-DEC-2004; 2004US-0636240P.
XX
PA (UYFL) UNIV FLORIDA RES FOUND INC.
XX
PI Chegini N, Luo X, Ding L, Williams RS;
XX
DR WPI; 2005-703565/72.
XX
PT Identifying a modulator of a gene that is differentially-expressed in
PT fibrotic tissue or during fibrogenesis, or a polypeptide encoded by the
PT gene, in a cell population by contacting the cell population with a test
PT agent.
XX
PS Disclosure; SEQ ID NO 723; 202pp; English.
XX
CC The invention describes a method of identifying a modulator of at least
CC one gene that is differentially-expressed in fibrotic tissue or during
CC fibrogenesis, or a polypeptide encoded by the differentially-expressed
CC gene, in a cell population, comprising contacting the cell population
CC with a test agent, and determining if the test agent modulates the
CC expression of the gene or biological activity of the polypeptide encoded
CC by the gene. Also described are: detecting a fibrotic disorder in a
CC subject; modulating gene expression in fibrotic tissue; and an array
CC comprising a substrate having addresses, where each address has a capture
CC probe that can specifically bind at least one polynucleotide that is
CC differentially expressed in fibrotic disorders, or its complement. The
CC method is useful in identifying a modulator of at least one gene that is
CC differentially-expressed in fibrotic tissue or during fibrogenesis, or a
CC polypeptide encoded by the differentially-expressed gene, in a cell
CC population for preparing a composition for diagnosing or treating
CC fibrotic disorders, e.g. uterine fibrosis. This sequence represents a
CC polynucleotide associated with detection and treatment of fibrotic
CC disorders. Note: This sequence does not appear in the printed
CC specification but has been obtained in electronic format directy from
CC WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 100998 BP; 25449 A; 22644 C; 23083 G; 29822 T; 0 U; 0 Other;

Query Match 37.5%; Score 150; DB 14; Length 100998;
Best Local Similarity 69.4%; Pred. No. 4.1e-31;
Matches 225; Conservative 1; Mismatches 81; Indels 17; Gaps 1;

```
Qy        30 GGAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTA 89
             | ||||| |    || ||| |       |    || |  |       |  |  ||||||||
Db     92984 GAAACCCCATCTCTACTAAAAGTACAAAAAATTAGCCGGGCGTGGTGGCGGGTGCCTGTA 93043

Qy        90 ATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCA 149
             |||||||||||| |||||||||||| ||| |||||||||||||||| ||| || ||||||
Db     93044 ATCCCAGCACTTGGGGAGGCCAAGGAGGGTGGATCACCTGAGGTCAGGAGTTCAAGACCA 93103

Qy       150 TCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGG 209
              |||||||||||||||||||||| |||| |||||||| ||||||||  |||  |||||||
Db     93104 GCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAGTACAAAAATCAGCCAGGCATGG 93163

Qy       210 TGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGC 269
             |||||  | |||||| |||||||||||| |||| | |||| |    ||| |   |   ||
Db     93164 TGGCAGGCGCCTGTAATCCCAGCTACTCGGGAGGCTGAGAATCACTTGAACCCTGGAGGC 93223

Qy       270 AGAGTGAGCCGAAATCACAGATCA----------------CAGAGTGAGCAGAGTGAGA 312
             ||||   || |  | |  ||||||                | | ||||  ||||||| |
Db     93224 AGAGGCTGCAGTGAGCCGAGATCACGCCACTGCACTCTAGCCTGGGTGATGAGAGTGAAA 93283

Qy       313 CKCCGTCTCAAAAACAACAACAAA 336
             |:|| |||||| ||||||||||||
Db     93284 CTCCATCTCAACAACAACAACAAA 93307
```

RESULT 31
AEB96535/c
ID AEB96535 standard; DNA; 151909 BP.
XX
AC AEB96535;
XX
DT 06-OCT-2005 (first entry)
XX
DE Human CABIN1 gene, SEQ ID 19.
XX
KW hepatitis C virus infection; antiinflammatory; hepatotropic; virucide;
KW liver cirrhosis; fibrosis; hepatoma; SNP detection; CABIN1; ds.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	variation	7588
FT		/*tag= a
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	10001
FT		/*tag= b
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	10054
FT		/*tag= c
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	15316
FT		/*tag= d
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	20223
FT		/*tag= e
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	21408
FT		/*tag= f
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	21444
FT		/*tag= g
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	28376
FT		/*tag= h
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	30258
FT		/*tag= i
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	30421
FT		/*tag= j
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	33211
FT		/*tag= k
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	34724
FT		/*tag= l
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	35061
FT		/*tag= m
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	39325
FT		/*tag= n
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	39373
FT		/*tag= o
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	39912
FT		/*tag= p
FT		/standard_name= "Single nucleotide polymorphism"

```
FT   variation      41614
FT                  /*tag= q
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      45740
FT                  /*tag= r
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      46172
FT                  /*tag= s
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      46553
FT                  /*tag= t
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      46771
FT                  /*tag= u
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      47848
FT                  /*tag= v
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      48021
FT                  /*tag= w
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      54446
FT                  /*tag= x
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      60910
FT                  /*tag= y
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      64350
FT                  /*tag= z
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      64676
FT                  /*tag= aa
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      70637
FT                  /*tag= ab
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      70818
FT                  /*tag= ac
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      71037
FT                  /*tag= ad
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      73664
FT                  /*tag= ae
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      74237
FT                  /*tag= af
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      75770
FT                  /*tag= ag
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      79208
FT                  /*tag= ah
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      80090
FT                  /*tag= ai
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      80839
FT                  /*tag= aj
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      81378
FT                  /*tag= ak
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      83040
FT                  /*tag= al
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      86161
FT                  /*tag= am
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      92229
FT                  /*tag= an
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      92791
FT                  /*tag= ao
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      95508
FT                  /*tag= ap
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      95692
FT                  /*tag= aq
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      96522
FT                  /*tag= ar
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      96770
FT                  /*tag= as
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      97080
FT                  /*tag= at
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      98461
FT                  /*tag= au
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      99705
FT                  /*tag= av
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      102545
FT                  /*tag= aw
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      103205
FT                  /*tag= ax
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      107881
FT                  /*tag= ay
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      109919
FT                  /*tag= az
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      110101
FT                  /*tag= ba
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      110228
FT                  /*tag= bb
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      110420
FT                  /*tag= bc
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      117399
FT                  /*tag= bd
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      117478
FT                  /*tag= be
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      117729
FT                  /*tag= bf
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      118132
FT                  /*tag= bg
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      120210
FT                  /*tag= bh
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      120523
FT                  /*tag= bi
FT                  /standard_name= "Single nucleotide polymorphism"
FT   variation      122339
FT                  /*tag= bj
FT                  /standard_name= "Single nucleotide polymorphism"

  Query Match            37.5%;  Score 149.8;  DB 14;  Length 151909;
  Best Local Similarity  71.0%;  Pred. No. 5.2e-31;
  Matches  211;  Conservative    1;  Mismatches   83;  Indels    2;  Gaps    1;

Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
              ||||||||||||||||| ||||||||||||||  ||||||||||||  |   |||
Db        587 CCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGTAGATCACCTGAGGTCGGGGATTCG 528
```

```
Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
              |||||| ||||||||||||||||||||||| ||| |||||||||||||||||||||||||
Db        527 AGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAAATACAAAAAATAGCTGG 468

Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG--TGAGCT 261
              |||||||||||     ||| || |||||||||||  |||  | |||   | ||  |    |
Db        467 GCATGGTGGCATGTGCCTATAATCCCAGCTACTTGGGAAGCTGAGGCAGGAGAATCGCTT 408

Qy        262 GAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTC 321
              ||    |    |      | |   |||      | | |  | |||||||||| |:| |||||
Db        407 GAACCTGGGAGGCAGAGCCACTGCACTCCAGCCTGGGGCAACAGAGTGAGGCTCTGTCTC 348

Qy        322 AAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAG 378
              |||||||  ||||||||  |||| ||  |  | | | |  ||    |  | | | ||
Db        347 CAAAACAAAAACAAAAACAAAAACAAAAAACAAAGAATAACCCAAAGGAGATGCAAG 291
```

RESULT 32
ACN44170
ID ACN44170 standard; DNA; 196686 BP.
XX
AC ACN44170;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genomic sequence hCG39530.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 484; 0pp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
SQ Sequence 196686 BP; 53978 A; 42758 C; 43862 G; 55372 T; 0 U; 716 Other;

```
  Query Match          37.5%;  Score 149.8;  DB 11;  Length 196686;
  Best Local Similarity   76.3%;  Pred. No. 5.7e-31;
  Matches  196;  Conservative    1;  Mismatches   58;  Indels    2;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| |||||||| ||||| ||||||||| ||||||||  |||
Db      29937 ACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGAGCGGATCACTTGAGGTCAGGAGT 29996

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              |||||||  |||||||||||||||| |||||||  |||||||||||||||| |||||  ||||
Db      29997 TCGAGACTAGCCTGGCCAACATGGGGAAACCCTGTCTTTACTAAAAATATAAAAATTAGC 30056

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||  |||||||||||||||||| ||| |||||||||||| | |||    | || |||
Db      30057 TGGGTGTGGTGGCACACACCTGTAATCCTAGCTACTCAGGAGGCTGAGGCAGGAGAGAGA 30116

Qy        261 T--GAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGT 318
              |  ||   ||| || | |    | |        |   |  | ||||||||||:|
Db      30117 TTGTAGTGAGCAAAGAGCGTGCCATTGCACTCCAGCCTGGGCAACAGAGTGAGACTCTAA 30176

Qy        319 CTCAAAAACAACAACAA 335
              |||||||| || || ||
Db      30177 CTCAAAAAAAAAAAAAA 30193
```

RESULT 33
ABL65836/c
ID ABL65836 standard; DNA; 5670 BP.
XX
AC ABL65836;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4173.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.

```
PR   02-OCT-2000; 2000US-0237173P.
PR   02-OCT-2000; 2000US-0237278P.
PR   02-OCT-2000; 2000US-0237294P.
PR   02-OCT-2000; 2000US-0237295P.
PR   02-OCT-2000; 2000US-0237316P.
PR   03-OCT-2000; 2000US-0237425P.
PR   03-OCT-2000; 2000US-0237598P.
PR   03-OCT-2000; 2000US-0237604P.
PR   03-OCT-2000; 2000US-0237606P.
PR   03-OCT-2000; 2000US-0237608P.
PR   01-NOV-2000; 2000US-0244867P.
PR   01-NOV-2000; 2000US-0245084P.
XX
PA   (AVAL-) AVALON PHARM.
XX
PI   Young PE,  Augustus M,  Carter KC,  Ebner R,  Endress G,  Horrigan S;
PI   Soppet DR,  Weaver Z;
XX
DR   WPI; 2002-188264/24.
XX
PT   Screening for anti-neoplastic agent involves exposing cells to a chemical
PT   agent to be tested for anti-neoplastic activity, and determining a change
PT   in expression of a gene of a signature gene set.
XX
PS   Claim 1; SEQ ID NO 4173; 44pp; English.
XX
CC   The present invention describes a method (M1) for screening for an anti-
CC   neoplastic agent. The method involves exposing cells to a chemical agent
CC   to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening an
CC   anti-neoplastic agent, and can be used for producing a product which is
CC   the data collected with respect to the anti-neoplastic agent as a result
CC   of M1, and the data is sufficient to convey the chemical structure and/or
CC   properties of the agent. M1 can be used in the treatment of cancer such
CC   as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC   prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC   cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC   cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC   tumour
XX
SQ   Sequence 5670 BP; 1267 A; 1630 C; 1616 G; 1157 T; 0 U; 0 Other;

  Query Match             37.4%;  Score 149.6;  DB 6;  Length 5670;
  Best Local Similarity   75.0%;  Pred. No. 2.2e-31;
  Matches  201;  Conservative    1;  Mismatches   60;  Indels    6;  Gaps    1;

Qy          83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
               |||||||||||||||||||| |||||||||||||| || ||||||||||||||||   | ||
Db        5307 GCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTAGGTGGATCACCTGAGGTCAGAGGTTC 5248

Qy         143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
               ||||||  |||||||||||||||||||||| || |||||||||||||||| ||||||
Db        5247 AAGACCAGTCTGGCCAACATGGTGAAACCCCGCCTCTACTAAAAATACAAAAATTAGCTG 5188

Qy         203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
               ||||||||||||  | |||||| ||||||||||  |||| | |||   | ||
Db        5187 GGCATGGTGGCAGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCT 5128

Qy         263 AGATCGCAGAGTGAGCCGAAATCACAGATC------ACAGAGTGAGCAGAGTGAGACKCC 316
               || | |||   |    |    ||  || |      ||   |  | |||||||||||:|
Db        5127 TGAACCCAGGAGGCAGAGGTTGCAGTGAGCACTCCAACCTGGGCAACAGAGTGAGACTCT 5068

Qy         317 GTCTCAAAAACAACAACAAAAAACAAAA 344
               |||||||||| || || |||||| ||||
Db        5067 GTCTCAAAAAAAAAAAAAAAAAAAAAAA 5040


RESULT 34
ACA64883/c
```

```
ID   ACA64883 standard; DNA; 5670 BP.
XX
AC   ACA64883;
XX
DT   27-JUN-2003  (first entry)
XX
DE   Human MB-1 gene (CD79a-B cell) DNA corresponding to U05259.
XX
KW   Human; chronic inflammatory joint disease; infection; tumour;
KW   antiinflammatory; cytostatic; antiarthritic; antirheumatic;
KW   immunosuppressive; gene therapy; etiological pathogenicity; ds.
XX
OS   Homo sapiens.
XX
PN   DE10127572-A1.
XX
PD   05-DEC-2002.
XX
PF   30-MAY-2001; 2001DE-01027572.
XX
PR   30-MAY-2001; 2001DE-01027572.
XX
PA   (PATH-) PATHOARRAY GMBH.
XX
PI   Haeupl T,  Ungethuem U,  Blaess S;
XX
DR   WPI; 2003-240797/24.
XX
PT   Reagents for diagnosis, study and therapy of chronic inflammatory joint
PT   and other diseases, comprises any of many specified genes or derived
PT   proteins.
XX
PS   Claim 1; Page; 12pp; German.
XX
CC   This invention describes a novel reagent for diagnosis, molecular
CC   definition and therapy of chronic inflammatory joint diseases, and other
CC   inflammatory disorders, infective or tumour diseases in humans. The
CC   products of the invention have antiinflammatory, cytostatic,
CC   antiarthritic, antirheumatic and immunosuppressive activity and can be
CC   used for gene therapy. The reagent of the invention and any proteins and
CC   antibodies derived from it, are used (i) for analysing tissue and blood
CC   samples for medical diagnosis; (ii) for diagnosis and characterisation of
CC   chronic joint diseases, on the basis of molecular characterisation, and
CC   determining the etiological pathogenicity principle of as yet
CC   uncharacterised inflammatory diseases, also monitoring progression and/or
CC   treatment of disease, and optimisation of therapy and (iii) for
CC   developing treatments for inflammatory diseases, particularly of joints,
CC   infections and tumours. ACA64801-ACA64965 represent human polynucleotides
CC   used in the method of the invention
XX
SQ   Sequence 5670 BP; 1267 A; 1630 C; 1616 G; 1157 T; 0 U; 0 Other;

  Query Match             37.4%;  Score 149.6;  DB 8;  Length 5670;
  Best Local Similarity   75.0%;  Pred. No. 2.2e-31;
  Matches  201;  Conservative    1;  Mismatches   60;  Indels    6;  Gaps    1;

Qy          83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
               |||||||||||||||||||| |||||||||||||| || ||||||||||||||||   | ||
Db        5307 GCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTAGGTGGATCACCTGAGGTCAGAGGTTC 5248

Qy         143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
               ||||||  |||||||||||||||||||||| || |||||||||||||||| ||||||
Db        5247 AAGACCAGTCTGGCCAACATGGTGAAACCCCGCCTCTACTAAAAATACAAAAATTAGCTG 5188

Qy         203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
               ||||||||||||  | |||||| ||||||||||  |||| | |||   | ||
Db        5187 GGCATGGTGGCAGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCT 5128

Qy         263 AGATCGCAGAGTGAGCCGAAATCACAGATC------ACAGAGTGAGCAGAGTGAGACKCC 316
               || | |||   |    |    ||  || |      ||   |  | |||||||||||:|
Db        5127 TGAACCCAGGAGGCAGAGGTTGCAGTGAGCACTCCAACCTGGGCAACAGAGTGAGACTCT 5068

Qy         317 GTCTCAAAAACAACAACAAAAAACAAAA 344
```

```
                ||||||||||  ||  ||  ||||||  ||||
Db       5067 GTCTCAAAAAAAAAAAAAAAAAAAAAAAA 5040

RESULT 35
AEF74508/c
ID   AEF74508 standard; DNA; 5670 BP.
XX
AC   AEF74508;
XX
DT   06-APR-2006  (first entry)
XX
DE   Human polynucleotide #22.
XX
KW   Diagnosis; gene regulation; gene expression;
KW   post traumatic stress disorder; psychiatric disorder; tranquilizer; gene;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO2006013561-A2.
XX
PD   09-FEB-2006.
XX
PF   02-AUG-2005; 2005WO-IL000824.
XX
PR   02-AUG-2004; 2004US-0592408P.
XX
PA   (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
PA   (HADA-) HADASIT MEDICAL RES SERVICES & DEV LTD.
XX
PI   Segman R,  Shalev A,  Goltser T,  Friedman N,  Shefi N,  Kaminski N;
XX
DR   WPI; 2006-145797/15.
XX
PT   New kit comprising 10 and no more than 574 polynucleotides capable of
PT   specifically binding at least one specific polynucleotide sequence,
PT   useful for determining predisposition of a subject to develop PTSD, or
PT   for diagnosing PTSD.
XX
PS   Claim 1; SEQ ID NO 22; 157pp; English.
XX
CC   The invention relates to a kit for determining predisposition of a
CC   subject to developing post-traumatic stress disorder (PTSD) comprising at
CC   least 10 and no more than 574 polynucleotides, where each of the
CC   polynucleotides is capable of specifically binding at least one specific
CC   polynucleotide sequence. The invention also relates to a kit for
CC   diagnosing PTSD in a subject, agents for the manufacture of the kits
CC   cited comprising the polynucleotides cited, and a microarray comprising
CC   at least 10 and no more than 904 oligonucleotides where each of the
CC   oligonucleotides is capable of specifically binding at least one specific
CC   polynucleotide sequence. The kit comprises each of the polynucleotides
CC   selected from an oligonucleotide molecule, a cDNA molecule, a genomic
CC   molecule and an RNA molecule. Each of the polynucleotides is at least 10
CC   and no more than 50 nucleic acids in length. Each of the polynucleotides
CC   is bound to a solid support. The kit also comprises at least one reagent
CC   suitable for detecting hybridization of the polynucleotides and at least
CC   one RNA transcript. The kit further comprises packaging materials
CC   packaging the at least one reagent and instructions for using the kit in
CC   determining predisposition of the subject to developing PTSD, or for
CC   diagnosing the disease. The microarray comprises oligonucleotides of at
CC   least 10 and no more than 40 nucleic acids in length. The agent is
CC   capable of regulating an expression level of at least one gene as a
CC   pharmaceutical or for the manufacture of a medicament identified for
CC   preventing PTSD. The kit is useful for determining predisposition of a
CC   subject to developing PTSD or for diagnosing PTSD. This sequence
CC   represents a human polynucleotide of the invention. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 5670 BP; 1267 A; 1630 C; 1616 G; 1157 T; 0 U; 0 Other;

  Query Match           37.4%;  Score 149.6;  DB 15;  Length 5670;
  Best Local Similarity  75.0%;  Pred. No. 2.2e-31;
  Matches  201;  Conservative    1;  Mismatches   60;  Indels    6;  Gaps    1;

Qy         83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
              ||||||||||||||||||| |||||||||||| || ||||||||||||||||    | ||
Db       5307 GCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTAGGTGGATCACCTGAGGTCAGAGGTTC 5248

Qy        143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
              ||||||  |||||||||||||||||||||| || ||||||||||||||| ||||||
Db       5247 AAGACCAGTCTGGCCAACATGGTGAAACCCCGCCTCTACTAAAAATACAAAAATTAGCTG 5188

Qy        203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
              ||||||||||||  | |||||| ||||||||||  |||| | |||   | ||
Db       5187 GGCATGGTGGCAGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCT 5128

Qy        263 AGATCGCAGAGTGAGCCGAAATCACAGATC------ACAGAGTGAGCAGAGTGAGACKCC 316
              || | |||   |    |    ||  || |      ||   |  | ||||||||||:|
Db       5127 TGAACCCAGGAGGCAGAGGTTGCAGTGAGCACTCCAACCTGGGCAACAGAGTGAGACTCT 5068

Qy        317 GTCTCAAAAACAACAACAAAAAACAAAA 344
              ||||||||||  ||  ||  ||||||  ||||
Db       5067 GTCTCAAAAAAAAAAAAAAAAAAAAAAA 5040


RESULT 36
AAK90749/c
ID   AAK90749 standard; DNA; 8205 BP.
XX
AC   AAK90749;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human digestive system antigen genomic sequence SEQ ID NO: 4325.
XX
KW   Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW   ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW   digestive system disorder; Meckel's diverticulum; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155314-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001324.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
```

PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-502630/55.
XX
PT Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognozing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX
PS Disclosure; SEQ ID NO 4325; 986pp; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a genomic DNA fragment
CC encoding a digestive system antigen of the invention
XX
SQ Sequence 8205 BP; 2216 A; 1646 C; 1763 G; 2580 T; 0 U; 0 Other;

Query Match 37.3%; Score 149; DB 4; Length 8205;
Best Local Similarity 75.5%; Pred. No. 3.6e-31;
Matches 200; Conservative 1; Mismatches 56; Indels 8; Gaps 1;

```
Qy        81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
             |||||||||||||||||||| || ||||||||| ||| ||||||||||||||| |||
Db      5600 ATGCCTGTAATCCCAGCACTTTGGAAGGCCAAGGCGGGTGGATCACCTGAGGTCAGGAGT 5541

Qy       141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
             ||||||||| |||||| |||||||||||||| |||| |||||||||||||||| ||||
Db      5540 TCGAGACCAGCCTGGCAAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGC 5481

Qy       201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
```

```
Db       5480 CGGGTGTGGTGGTGTGCGCCTATAATCCCAGCTACTCAGGAGGCAGAGGTTGCAGTGAGC 5421

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320

Db       5420 CAAGACTGC--------GCCATTACACTCCGGCCTGGGCAACAAGAGCAAAACTCCATCT 5369

Qy        321 CAAAAACAACAACAAAAAACAAAAA 345

Db       5368 CAAAAAAAAAAGAAAAAAAAAGACA 5344
```

RESULT 37
AAI62948/c
ID AAI62948 standard; DNA; 8205 BP.
XX
AC AAI62948;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human genomic DNA SEQ ID NO 276.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200155449-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001346.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUL-2000; 2000US-0216880P.
PR 14-JUL-2000; 2000US-0218290P.
PR 14-AUG-2000; 2000US-0225447P.
PR 01-SEP-2000; 2000US-0229343P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0231243P.
PR 25-SEP-2000; 2000US-0234997P.
PR 29-SEP-2000; 2000US-0236367P.
PR 13-OCT-2000; 2000US-0239937P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246528P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249265P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-476225/51.
XX
PT Novel plasma membrane associated proteins useful for diagnosing,
PT treating, preventing and/or prognosing disorders related to the proteins,
PT including cancer, immune response and neuronal disorders.
XX
PS Example 2; SEQ ID NO 276; 532pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAI62752-AAI62961) and proteins
CC (AAM42347-AAM42415) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 8205 BP; 2216 A; 1646 C; 1763 G; 2580 T; 0 U; 0 Other;

Query Match 37.3%; Score 149; DB 4; Length 8205;
Best Local Similarity 75.5%; Pred. No. 3.6e-31;
Matches 200; Conservative 1; Mismatches 56; Indels 8; Gaps 1;

```
Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140

Db       5600 ATGCCTGTAATCCCAGCACTTTGGAAGGCCAAGGCGGGTGGATCACCTGAGGTCAGGAGT 5541

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200

Db       5540 TCGAGACCAGCCTGGCAAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCC 5481

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260

Db       5480 CGGGTGTGGTGGTGTGCGCCTATAATCCCAGCTACTCAGGAGGCAGAGGTTGCAGTGAGC 5421

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320

Db       5420 CAAGACTGC--------GCCATTACACTCCGGCCTGGGCAACAAGAGCAAAACTCCATCT 5369

Qy        321 CAAAAACAACAACAAAAAACAAAAA 345

Db       5368 CAAAAAAAAAAGAAAAAAAAAGACA 5344
```

RESULT 38
ADO79404
ID ADO79404 standard; DNA; 89900 BP.
XX
AC ADO79404;
XX
DT 26-AUG-2004 (first entry)
XX
DE DPF3 region, SEQ ID 3.
XX
KW Cytostatic; Gene therapy; breast cancer; human; DLG1; KIAA0783; DPF3;
KW CENPC1; gene; ds; SNP; single nucleotide polymorphism;
KW D4, zinc and double PHD fingers, family 3; CERD4; cer-d4; FLJ14079;
KW 2810403B03Rik; Rho family guanine-nucleotide exchange factor;
KW chromosome 14q24.3-q31.1.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation 160

```
FT                   /*tag= a
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a A/C SNP"
FT   variation       6053
FT                   /*tag= b
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a T/G SNP"
FT   variation       9719
FT                   /*tag= c
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a A/G SNP"
FT   variation       10481
FT                   /*tag= d
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a T/C SNP"
FT   variation       10676
FT                   /*tag= e
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a A/T SNP"
FT   variation       17179
FT                   /*tag= f
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a C/G SNP"
FT   variation       18561
FT                   /*tag= g
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a A/T SNP"
FT   variation       18658
FT                   /*tag= h
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a G/C SNP"
FT   variation       18694
FT                   /*tag= i
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a A/G SNP"
FT   variation       18858
FT                   /*tag= j
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a T/C SNP"
FT   variation       24582
FT                   /*tag= k
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a G/A SNP"
FT   variation       24683
FT                   /*tag= l
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a G/A SNP"
FT   variation       24767
FT                   /*tag= m
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a T/C SNP"
FT   variation       27402
FT                   /*tag= n
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a A/G SNP"
FT   variation       28150
FT                   /*tag= o
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a T/G SNP"
FT   variation       28494
FT                   /*tag= p
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a T/C SNP"
FT   variation       32003
FT                   /*tag= q
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a A/C SNP"
FT   variation       35588
FT                   /*tag= r
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a T/C SNP"
FT   variation       35619
FT                   /*tag= s
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a T/C SNP"
FT   variation       35856
FT                   /*tag= t
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a G/A SNP"
FT   variation       36254
FT                   /*tag= u
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a G/C SNP"
FT   variation       37314
FT                   /*tag= v
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a G/A SNP"
FT   variation       40033
FT                   /*tag= w
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a T/G SNP"
FT   variation       40095
FT                   /*tag= x
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a G/C SNP"
FT   variation       42593
FT                   /*tag= y
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a A/C SNP"
FT   variation       42799
FT                   /*tag= z
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a A/G SNP"
FT   variation       43090
FT                   /*tag= aa
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a G/A SNP"
FT   variation       46683
FT                   /*tag= ab
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a A/G SNP"
FT   variation       49774
FT                   /*tag= ac
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a A/G SNP"
FT   variation       51796
FT                   /*tag= ad
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a T/C SNP"
FT   variation       52079
FT                   /*tag= ae
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a A/T SNP"
FT   variation       53857
FT                   /*tag= af
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a T/C SNP"
FT   variation       53971
FT                   /*tag= ag
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a A/C SNP"
FT   variation       55899
FT                   /*tag= ah
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a T/C SNP"
FT   variation       60682
FT                   /*tag= ai
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a G/A SNP"
FT   variation       61291
FT                   /*tag= aj
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a T/C SNP"
FT   variation       72720
FT                   /*tag= ak
FT                   /standard_name= "Single nucleotide polymorphism"
```

```
FT                   /note= "This SNP is described as a G/A SNP"
FT   variation       72752
FT                   /*tag=  al
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a A/C SNP"
FT   variation       85507
FT                   /*tag=  am
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a A/G SNP"
FT   variation       89751
FT                   /*tag=  an
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "This SNP is described as a T/A SNP"
XX
PN   WO2004047514-A2.
XX
PD   10-JUN-2004.
XX
PF   25-NOV-2003; 2003WO-US037943.
XX
PR   25-NOV-2002; 2002US-0429136P.
PR   24-JUL-2003; 2003US-0490234P.
XX
PA   (SEQU-) SEQUENOM INC.
XX
PI   Roth RB,  Nelson MR,  Braun A,  Kammerer SM,  Reneland R;
XX
DR   WPI; 2004-441037/41.
XX
PT   Identifying a subject at risk of breast cancer by detecting the presence
PT   of polymorphic variations in the DLG1, KIAA0783, DPF3 or CENPC1 regions
PT   which are associated with breast cancer in a nucleic acid sample from a
PT   subject.
XX
PS   Claim 24; Fig 3; 227pp; English.
XX

  Query Match             37.3%;  Score 149;  DB 12;  Length 89900;
  Best Local Similarity   76.4%;  Pred. No. 7.5e-31;
  Matches  207;  Conservative    1;  Mismatches   61;  Indels    2;  Gaps    2;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||||| |||||||  ||||||| |||||| |||||||||| |||
Db       3670 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGTGGATCGCCTGAGGTCAGGAGT 3729

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||||||||||||||| ||||||||| |||||||||||||||| ||||
Db       3730 TCGAGACCAGCCTGGCCAACATGGTGACACCCCGTCTCTACTAAAAATACAAAAATTAGC 3789

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCG-GAGATTGCAGTGAG 259
              ||||| |||||||| || |||||| |||||||||||||| ||| ||    ||   |||    |
Db       3790 TGGGCCTGGTGGCAGACGCCTGTAATCCCAGCTACTCAAGAGTCGCTTGAACTCAGAAGG 3849

Qy        260 CTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAG-CAGAGTGAGACKCCGT 318
              || | |      |      |  |   ||   ||    | | | ||| |||||||:| ||
Db       3850 TTGTGGTGAGTTGGGATTGTGCCACTGCACTCCAGCCTGGGCGACAGGGTGAGACTCTGT 3909

Qy        319 CTCAAAAACAACAACAAAAAACAAAAAAACC 349
              || ||||| || || |||||| |||||||||
Db       3910 CTAAAAAAAAAAAAAAAAAAAAAAAAAAAACC 3940


RESULT 39
AEF92655/c
ID   AEF92655 standard; cDNA; 174318 BP.
XX
AC   AEF92655;
XX
DT   20-APR-2006  (first entry)
XX
DE   Human Ras effector protein Rin3, cDNA.
XX
KW   Diagnosis; ss; gene; neurodegenerative disease; Alzheimers disease;
KW   dementia; cognitive disorder; Parkinsons disease; microarray;
KW   neuroprotective; nootropic; antiparkinsonian; screening; prognosis.
XX
OS   Homo sapiens.
XX
PN   WO2006020269-A2.
XX
PD   23-FEB-2006.
XX
PF   19-JUL-2005; 2005WO-US025491.
XX
PR   19-JUL-2004; 2004US-0589318P.
XX
PA   (UYRP ) UNIV ROCHESTER.
XX
PI   Coleman PD,  Federoff HJ,  Maguire-Zeiss K,  Mhyre TR,  Kurlan RM;
PI   Cox C,  Marshall F;
XX
DR   WPI; 2006-184393/19.
XX
PT   Use of biomarkers for neurodegenerative disease for, e.g. diagnosing
PT   neurodegenerative disease, screening therapeutic agent for treating
PT   neurodegenerative disease, or monitoring neurodegenerative disease
PT   progression.
XX
PS   Example 1; SEQ ID NO 82; 552pp; English.
XX
CC   The invention relates to using a biomarker for a neurodegenerative
CC   disease for diagnosing a neurodegenerative disease, screening a
CC   therapeutic agent for treating a neurodegenerative disease, monitoring a
CC   neurodegenerative disease progression, monitoring a response to a
CC   neurodegenerative disease treatment, identifying a risk for a
CC   neurodegenerative disease, and differentially diagnosing a
CC   neurodegenerative disease in a test subject. Also included are diagnosing
CC   a neurodegenerative disease in a subject, screening for a therapeutic
CC   agent for the treatment of a neurodegenerative disease, monitoring a
CC   neurodegenerative disease progression in a subject, monitoring a response
CC   to a neurodegenerative disease treatment in a subject, identifying a risk
CC   for a neurodegenerative disease in a test subject, differentially
CC   diagnosing a neurodegenerative disease in a test subject, a solid support
CC   (comprising one or more biomarkers, where the biomarker is one or more
CC   proteins comprising HSP60, Dihydrolipoamide dehydrogenase, ER-60
CC   protease, Glucose-6-phosphate dehydrogenase, ATP-synthase beta chain,
CC   Annexin I, 14-3-3 epsilon, Prohibitin, Phosphoglycerate mutase 1,
CC   Apolipoprotein AI, Superoxide dismutase, RNA-binding protein regulatory
CC   subunit, Chain A thioredoxin peroxidase B, RAS-related protein RAP1B,
CC   Tumor rejection antigen, Haptoglobin, Fibrin beta, or its combinations)
CC   and a solid support comprising one or more biomarkers (where the
CC   biomarker is one or more transcripts comprising cyclin D1, cyclin B,
CC   cyclin G1, wee1, hTR2, CDC25b, GSK3 beta, protein kinase C alpha, C5, C1
CC   inhibitor, IL-17r, IL-8, LIF, TNF-alpha, IL-10r, Alpha-1
CC   antichymotrypsin, HSP 27, HSP 90, crystalline, GAPDH, ferritin H,
CC   ferritin L, cox 1, cox 2, transferrin, or its combinations). The
CC   biomarkers for the neurodegenerative disease are useful for diagnosing a
CC   neurodegenerative disease, screening a therapeutic agent for treating a
CC   neurodegenerative disease, monitoring a neurodegenerative disease
CC   progression monitoring a response to a neurodegenerative disease
CC   treatment, identifying a risk for a neurodegenerative disease, and
CC   differentially diagnosing a neurodegenerative disease, e.g. Alzheimer's
CC   disease and Parkinson's disease, in a test subject. The present sequence
CC   is a cDNA for a human biomarker, used in a microarray in the method of
CC   the invention. NOTE: The specification describes AEF92574-AEF92813 (table
CC   4) as cDNA sequences yet some are protein sequences and some genomic DNA.
XX
SQ   Sequence 174318 BP; 43826 A; 41095 C; 42605 G; 46792 T; 0 U; 0 Other;

  Query Match             37.3%;  Score 149;  DB 15;  Length 174318;
  Best Local Similarity   73.5%;  Pred. No. 9.2e-31;
  Matches  202;  Conservative    1;  Mismatches   71;  Indels    1;  Gaps    1;

Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
              |||||||| ||||||||| || ||||| ||||||| |||||||||||||||| |||    |
Db      82961 CCTGTAATGCCAGCACTTTGGAAGGCCGAGGTGGGAGGATCACCTGAGGTCAGGAGTCTG 82902
```

```
Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
              |||||| |||||||||||||||||||||| |||| | |||||||||||||| ||||  |
Db      82901 AGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTGCTAAAAATACAAAAATTAGCCAG 82842

Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
              || |||||||||||| |||||||||||||||||||  |||| | |||   | ||
Db      82841 GCGTGGTGGCACACGCCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCTT 82782

Qy        264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAG-CAGAGTGAGACKCCGTCTCA 322
              || | |||   | | |    ||  || | ||| || | ||  |  ||||:| ||||||
Db      82781 GAACCCAGGAGGTGGAGGTTGCAGTGAGCTGAGATTGCGCCACTGCAAGACTCTGTCTCA 82722

Qy        323 AAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
                || || || ||||||  |||| |  | || | |
Db      82721 CCAAAAAAAAAAAAAAAAGAAAAGAAAAGAAAAA 82687


RESULT 40
ADX98573
ID   ADX98573 standard; DNA; 285300 BP.
XX
AC   ADX98573;
XX
DT   05-MAY-2005  (first entry)
XX
DE   Human D4, zinc and double PHD fingers, family 3 (DPF3) genomic DNA.
XX
KW   SNP detection; breast tumor; endocrine disease;
KW   gynecology and obstetrics; neoplasm; cytostatic; metastasis;
KW   gene therapy; RNA interference; chromosome 14; ds; SNP;
KW   single nucleotide polymorphism;
KW   D4, zinc and double PHD fingers, family 3; DPF3;
KW   guanine-nucleotide exchange factor.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       207
FT                   /*tag= a
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       486
FT                   /*tag= b
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       1745
FT                   /*tag= c
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       1922
FT                   /*tag= d
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       2190
FT                   /*tag= e
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       2590
FT                   /*tag= f
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       2637
FT                   /*tag= g
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       2804
FT                   /*tag= h
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       2806
FT                   /*tag= i
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       2895
FT                   /*tag= j
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       3109
FT                   /*tag= k
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       3185
FT                   /*tag= l
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       3355
FT                   /*tag= m
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       3642
FT                   /*tag= n
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       3805
FT                   /*tag= o
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       4236
FT                   /*tag= p
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       4331
FT                   /*tag= q
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       4509
FT                   /*tag= r
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       4959
FT                   /*tag= s
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       5009
FT                   /*tag= t
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       5676
FT                   /*tag= u
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       6507
FT                   /*tag= v
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       6695
FT                   /*tag= w
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       6717
FT                   /*tag= x
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       7131
FT                   /*tag= y
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       7873
FT                   /*tag= z
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       7922
FT                   /*tag= aa
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       8652
FT                   /*tag= ab
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       9817
FT                   /*tag= ac
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       10272
FT                   /*tag= ad
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       10823
FT                   /*tag= ae
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       11465
FT                   /*tag= af
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       11639
FT                   /*tag= ag
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       12177
FT                   /*tag= ah
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       12604
FT                   /*tag= ai
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       13363
FT                   /*tag= aj
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
```

```
FT   variation       13454
FT                   /*tag= ak
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       13540
FT                   /*tag= al
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       13923
FT                   /*tag= am
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       14507
FT                   /*tag= an
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       16550
FT                   /*tag= ao
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       17641
FT                   /*tag= ap
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       18903
FT                   /*tag= aq
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       19395
FT                   /*tag= ar
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       19527
FT                   /*tag= as
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       20369
FT                   /*tag= at
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       20505
FT                   /*tag= au
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       20562
FT                   /*tag= av
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       20907
FT                   /*tag= aw
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       20949
FT                   /*tag= ax
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       21278
FT                   /*tag= ay
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       21314
FT                   /*tag= az
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       21905
FT                   /*tag= ba
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       22252
FT                   /*tag= bb
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       22941
FT                   /*tag= bc
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       23542
FT                   /*tag= bd
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       24677
FT                   /*tag= be
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       25009
FT                   /*tag= bf
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       25618
FT                   /*tag= bg
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       26082
FT                   /*tag= bh
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   variation       26136
FT                   /*tag= bi

  Query Match             37.3%;  Score 149;  DB 14;  Length 285300;
  Best Local Similarity   76.4%;  Pred. No. 1.1e-30;
  Matches  207;  Conservative    1;  Mismatches   61;  Indels    2;  Gaps    2;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              ||||||||||||||||||||| |||||||  |||||||  |||||  |||||||||| |||
Db     164620 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGTGGATCGCCTGAGGTCAGGAGT 164679

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||||||||||||||| ||||||||| |||||||||||||||| ||||
Db     164680 TCGAGACCAGCCTGGCCAACATGGTGACACCCCGTCTCTACTAAAAATACAAAAATTAGC 164739

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCG-GAGATTGCAGTGAG 259
              ||||| |||||||| || |||||| |||||||||||| ||| ||   ||   |||    |
Db     164740 TGGGCCTGGTGGCAGACGCCTGTAATCCCAGCTACTCAAGAGTCGCTTGAACTCAGAAGG 164799

Qy        260 CTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAG-CAGAGTGAGACKCCGT 318
              || | |      |      | |   ||   ||    | | | ||| |||||||:| ||
Db     164800 TTGTGGTGAGTTGGGATTGTGCCACTGCACTCCAGCCTGGGCGACAGGGTGAGACTCTGT 164859

Qy        319 CTCAAAAACAACAACAAAAAACAAAAAAACC 349
              || ||||| || || |||||| ||||||||
Db     164860 CTAAAAAAAAAAAAAAAAAAAAAAAAAACC 164890


RESULT 41
ABK83568/c
ID   ABK83568 standard; DNA; 201143 BP.
XX
AC   ABK83568;
XX
DT   29-AUG-2002  (first entry)
XX
DE   Human DNA differentially expressed in granulocytic cells #139.
XX
KW   Human; ds; granulocytic cell; DNA chip; bacterial infection;
KW   viral infection; parasitic infection; protozoal infection;
KW   fungal infection; sterile inflammatory disease; psoriasis;
KW   rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW   cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW   adult respiratory distress syndrome; inflammatory bowel disease;
KW   Crohn's disease; ulcerative colitis; periodontal disease;
KW   granulocyte activation; chronic inflammation; allergy.
XX
OS   Homo sapiens.
XX
PN   WO200228999-A2.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US030821.
XX
PR   03-OCT-2000; 2000US-0237189P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Beazer-Barclay Y,  Weissman SM,  Yamaga S,  Vockley J;
XX
DR   WPI; 2002-435328/46.
XX
PT   Detecting granulocyte activation by detecting differential expression of
PT   genes associated with granulocyte activation, which serves as diagnostic
PT   markers that is useful for monitoring disease states and drug toxicity.
XX
PS   Claim 1; SEQ ID NO 139; 114pp; English.
XX
CC   The invention relates to detecting (M1) granulocyte (GC) activation
CC   (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC   DNA chip analysis as given in the specification, and comparing the
CC   expression level to an expression level in an unactivated GC, where
CC   differential expression of Gs is indicative of GCA. Also included are
```

```
CC   modulating (M2) GA by contacting GC with an agent that alters the
CC   expression of at least one gene in Gs; (2) screening (M3) for an agent
CC   capable of modulating GCA or an inflammation (especially chronic) in a
CC   tissue, an allergic response in a subject, exposure of a subject to a
CC   pathogen or sterile inflammatory disease using the gene expression
CC   profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC   tissue, an allergic response in a subject, exposure of a subject to a
CC   pathogen or sterile inflammatory disease, by detecting the level of
CC   expression in a sample of the tissue of gene(s) from Gs, where the level
CC   of expression of the gene is indicative of inflammation; (4) treating
CC   (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC   response in a subject, exposure of a subject to a pathogen or sterile
CC   inflammatory disease, by contacting a tissue having inflammation with an
CC   agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC   is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC   for screening an agent capable of modulating GCA preferably in an
CC   inflammation in a tissue; M4 is useful for detecting an inflammation
CC   (especially chronic) in a tissue, an allergic response in a subject,
CC   exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC   psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC   cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC   respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC   disease, ulcerative colitis, periodontal disease; also bacterial
CC   infection, viral infection, parasitic infection, protozoal infection,
CC   fungal infection and M5 is useful for treating one of the above
CC   conditions. The present sequence represents a gene differentially
CC   expressed in granulocytes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 201143 BP; 50606 A; 49308 C; 49683 G; 51546 T; 0 U; 0 Other;

  Query Match             37.2%;  Score 148.8;  DB 6;  Length 201143;
  Best Local Similarity   70.7%;  Pred. No. 1.1e-30;
  Matches  229;  Conservative    1;  Mismatches   83;  Indels   11;  Gaps    2;

Qy          37 AAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAG 96
               | || |||| || |     ||||||||||||||      |    ||||||||||||||||
Db        3229 ACATTTTAACAATATTAATTCAGGCCAGGCATG----GTGGCTCATGCCTGTAATCCCAG 3174

Qy          97 CACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGC 156
               |||||  ||||||  ||   ||| ||||||||||||||| ||| |||||||||| ||||||
Db        3173 CACTTTTGGAGGCTGAGACTGGCAGATCACCTGAGGTCAGGAGTTCGAGACCAGCCTGGC 3114

Qy         157 CAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACA 216
               |||||||||||||||| ||| ||||||||||||||||| ||||  ||||||||||||
Db        3113 CAACATGGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGCCAGGCATGGTGGCAGG 3054

Qy         217 CACCTGTAGTCCCAGCTACTCAGGA-------GCCGGAGATTGCAGTGAGCTGAGATCGC 269
               |||||||| ||||||||||||||||       || ||||| |     ||| |    |   ||
Db        3053 CACCTGTAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATCGCTTGAACCCTGGAGGC 2994

Qy         270 AGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAA 329
               ||||  || | || |   || |  || ||     |||||  | ||:| |||  ||||| ||
Db        2993 GGAGTTTGCAGCAAGCTGAGTTGGCATAGGCGACAGAGCAAAACTCTGTCAAAAAAAAAA 2934

Qy         330 CAACAAAAAACAAAAAAACCATAA 353
                 |      |  ||  ||||| ||
Db        2933 AATTTGCTGAGTAATGAACCAAAA 2910


RESULT 42
ADA02774/c
ID   ADA02774 standard; DNA; 44075 BP.
XX
AC   ADA02774;
XX
DT   06-NOV-2003 (first entry)
XX
DE   Human ARHGEF1 carcinoma associated gene, SEQ ID NO:1292.
XX
KW   Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW   prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003057146-A2.
XX
PD   17-JUL-2003.
XX
PF   26-DEC-2002; 2002WO-US041414.
XX
PR   26-DEC-2001; 2001US-00035832.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW;
XX
DR   WPI; 2003-587068/55.
XX
PT   New recombinant nucleic acid encoding carcinoma associated protein,
PT   useful for preparing compositions for treating carcinomas.
XX
PS   Claim 1; SEQ ID NO 1292; 245pp; English.
XX
CC   The invention relates to recombinant carcinoma associated (CA) nucleic
CC   acid sequences from mouse and human (ADA01482-ADA03094), and to
CC   recombinant carcinoma associated proteins (CAP) encoded by them. The
CC   invention also encompasses expression vectors and host cells comprising a
CC   CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC   binds to the protein, and a biochip comprising CA nucleic acid or
CC   fragments thereof. The sequences of the invention were identified using
CC   oncogenic retroviruses, which insert into the genome of the host organism
CC   at random. Many of these do not carry transduced host oncogenes or
CC   pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC   direct consequence of the effects of proviral integration into host
CC   protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC   carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC   leukaemia) or a propensity to carcinoma by determination of the sequence
CC   of a CA gene, or by determination of CA gene expression in particular
CC   tissues. CA nucleic acids, proteins and antibodies are also useful as
CC   therapeutic agents and in screening and evaluating drug candidates. The
CC   present sequence represents a specifically claimed human CA nucleic acid
CC   sequence of the invention. Note: The complete sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 44075 BP; 8314 A; 11402 C; 11137 G; 8371 T; 0 U; 4851 Other;

  Query Match             37.2%;  Score 148.6;  DB 9;  Length 44075;
  Best Local Similarity   74.9%;  Pred. No. 7.7e-31;
  Matches  200;  Conservative    1;  Mismatches   60;  Indels    6;  Gaps    1;

Qy          83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
               |||||||||||||||||| |||||||||||| || ||||||||||||||||    | ||
Db        8665 GCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTAGGTGGATCACCTGAGGTCAGAGGTTC 8606

Qy         143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
               ||||||  |||||||||||||||||||||| || |||||||||||||||| ||||||
Db        8605 AAGACCAGTCTGGCCAACATGGTGAAACCCCGCCTCTACTAAAAATACAAAAATTAGCTG 8546

Qy         203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
               |||||||||||  | |||||| ||||||||||  |||| | |||   | ||
Db        8545 GGCATGGTGGCAGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCT 8486

Qy         263 AGATCGCAGAGTGAGCCGAAATCACAGATC------ACAGAGTGAGCAGAGTGAGACKCC 316
               || | |||   |     |    ||  || |       ||   |  | |||||||||||:|
Db        8485 TGAACCCAGGAGGCAGAGGTTGCAGTGAGCACTCCAACCTGGGCAACAGAGTGAGACTCT 8426

Qy         317 GTCTCAAAAACAACAACAAAAAACAAA 343
               |||||||||| || || ||||||| |||
Db        8425 GTCTCAAAAAAAAAAAAAAAAAAAAAA 8399
```

```
RESULT 43
ADB72512/c
ID   ADB72512 standard; DNA; 44075 BP.
XX
AC   ADB72512;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human ARHGEF1 gene.
XX
KW   human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW   cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS   Homo sapiens.
XX
PN   WO2003008583-A2.
XX
PD   30-JAN-2003.
XX
PF   26-DEC-2001; 2001WO-US051291.
XX
PR   02-MAR-2001; 2001US-00798586.
PR   23-OCT-2001; 2001US-00004113.
PR   08-NOV-2001; 2001US-00052482.
PR   30-NOV-2001; 2001US-00997722.
PR   20-DEC-2001; 2001US-00034650.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW,  Engelhard EK;
XX
DR   WPI; 2003-239337/23.
XX
PT   New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT   cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
PS   Claim 1; SEQ ID NO 340; 2304pp; English.
XX
CC   The invention relates to a novel recombinant nucleic acid comprising a
CC   nucleotide sequence selected from any of the 660 sequences fully defined
CC   in the specification. A polynucleotide of the invention has cytostatic
CC   activity, and may have a use in gene therapy, or in a vaccine. The
CC   recombinant nucleic acids and polypeptides are useful for treating
CC   carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC   sarcomas. The present sequence represents a human gene of the invention.
XX
SQ   Sequence 44075 BP; 8314 A; 11402 C; 11137 G; 8371 T; 0 U; 4851 Other;

  Query Match             37.2%;  Score 148.6;  DB 10;  Length 44075;
  Best Local Similarity   74.9%;  Pred. No. 7.7e-31;
  Matches  200;  Conservative    1;  Mismatches   60;  Indels    6;  Gaps    1;

Qy         83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
              |||||||||||||||||| |||||||||||| || ||||||||||||||||   | ||
Db       8665 GCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTAGGTGGATCACCTGAGGTCAGAGGTTC 8606

Qy        143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
              |||||| |||||||||||||||||||||| || |||||||||||||||| ||||||
Db       8605 AAGACCAGTCTGGCCAACATGGTGAAACCCCGCCTCTACTAAAAATACAAAAATTAGCTG 8546

Qy        203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
              |||||||||||  | |||||| ||||||||||  |||| | |||   | ||
Db       8545 GGCATGGTGGCAGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCT 8486

Qy        263 AGATCGCAGAGTGAGCCGAAATCACAGATC------ACAGAGTGAGCAGAGTGAGACKCC 316
              || | |||   |    |    ||  || |      ||   |  | ||||||||||:|
Db       8485 TGAACCCAGGAGGCAGAGGTTGCAGTGAGCACTCCAACCTGGGCAACAGAGTGAGACTCT 8426

Qy        317 GTCTCAAAAACAACAACAAAAAACAAA 343
              |||||||||| || || ||||||| |||
Db       8425 GTCTCAAAAAAAAAAAAAAAAAAAAA 8399


RESULT 44
ADC85254/c
ID   ADC85254 standard; DNA; 44075 BP.
XX
AC   ADC85254;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Human Arhgef1 genomic sequence.
XX
KW   Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
KW   secreted; transmembrane; intracellular; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003045230-A2.
XX
PD   05-JUN-2003.
XX
PF   02-DEC-2002; 2002WO-US038582.
XX
PR   30-NOV-2001; 2001US-00997722.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW,  Engelhard EK;
XX
DR   WPI; 2003-513603/48.
XX
PT   New recombinant nucleic acid comprising a nucleotide sequence of any of
PT   the carcinoma-associated (CA) genes, useful for screening for drug
PT   candidates for diagnosing or treating carcinomas.
XX
PS   Claim 1; SEQ ID NO 40; 983pp; English.
XX
CC   The invention relates to a recombinant nucleic acid comprising a
CC   nucleotide sequence selected from any of the fully defined carcinoma-
CC   associated (CA) genes from the 50 tables given in the specification. The
CC   CA proteins are secreted, transmembrane or intracellular proteins. The
CC   recombinant nucleic acids are useful for screening for drug candidates
CC   for diagnosing or treating carcinomas. Sequences given in ADC85215-
CC   ADC85514 represent CA genes of the invention.
XX
SQ   Sequence 44075 BP; 8314 A; 11401 C; 11138 G; 8371 T; 0 U; 4851 Other;

  Query Match             37.2%;  Score 148.6;  DB 10;  Length 44075;
  Best Local Similarity   74.9%;  Pred. No. 7.7e-31;
  Matches  200;  Conservative    1;  Mismatches   60;  Indels    6;  Gaps    1;

Qy         83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
              |||||||||||||||||| |||||||||||| || ||||||||||||||||   | ||
Db       8665 GCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTAGGTGGATCACCTGAGGTCAGAGGTTC 8606

Qy        143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
              |||||| |||||||||||||||||||||| || |||||||||||||||| ||||||
Db       8605 AAGACCAGTCTGGCCAACATGGTGAAACCCCGCCTCTACTAAAAATACAAAAATTAGCTG 8546

Qy        203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
              |||||||||||  | |||||| ||||||||||  |||| | |||   | ||
Db       8545 GGCATGGTGGCAGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCT 8486

Qy        263 AGATCGCAGAGTGAGCCGAAATCACAGATC------ACAGAGTGAGCAGAGTGAGACKCC 316
              || | |||   |    |    ||  || |      ||   |  | ||||||||||:|
Db       8485 TGAACCCAGGAGGCAGAGGTTGCAGTGAGCACTCCAACCTGGGCAACAGAGTGAGACTCT 8426

Qy        317 GTCTCAAAAACAACAACAAAAAACAAA 343
              |||||||||| || || ||||||| |||
Db       8425 GTCTCAAAAAAAAAAAAAAAAAAAAA 8399


RESULT 45
ADM74369/c
```

ID ADM74369 standard; DNA; 44075 BP.
XX
AC ADM74369;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human carcinoma associated (CA) nucleic acid #19.
XX
KW Human; carcinoma associated nucleic acid; CA nucleic acid; gene; ds;
KW carcinoma associated protein; CAP; carcinoma; leukaemia; lymphoma;
KW cytostatic.
XX
OS Homo sapiens.
XX
PN US2004072154-A1.
XX
PD 15-APR-2004.
XX
PF 30-NOV-2001; 2001US-00997722.
XX
PR 22-DEC-2000; 2000US-00747377.
PR 02-MAR-2001; 2001US-00798586.
XX
PA (MORR/) MORRIS D W.
PA (ENGE/) ENGELHARD E K.
XX
PI Morris DW, Engelhard EK;
XX
DR WPI; 2004-328562/30.
XX
PT New carcinoma associated gene or protein, useful for preparing a
PT composition for diagnosing or treating carcinoma e.g., leukemia or
PT lymphoma.
XX
PS Claim 1; SEQ ID NO 40; 29pp; English.
XX
CC The invention relates to new recombinant nucleic acids. The invention
CC also relates to a host cell comprising a recombinant nucleic acid or
CC expression vector, an expression vector comprising a recombinant nucleic
CC acid, a recombinant protein, a method of screening for drug candidates, a
CC method of screening for a bioactive agent capable of binding to a
CC carcinoma associated protein (CAP) encoded by a nucleotide sequence, a
CC method of screening for a bioactive agent capable of modulating the
CC activity of a CAP, a method of evaluating the effect of a candidate
CC carcinoma drug, a method of diagnosing carcinoma, a method for inhibiting
CC the activity of a CAP, a method of treating carcinomas, a method of
CC neutralising the effect of a CAP and a method of diagnosing carcinoma or
CC propensity to carcinoma. A method of evaluating the effect of a candidate
CC carcinoma drug comprises administering the drug to a patient, removing a
CC cell sample from the patient and determining alterations in the
CC expression or activation of a gene comprising the nucleotide sequence. A
CC method of diagnosing carcinoma comprises determining the expression of
CC one or more genes comprising the nucleic acid sequence in a first tissue
CC type of a first individual and comparing the expression of the gene from
CC a second normal tissue type from the first individual or a second
CC unaffected individual, where a difference in the expression indicates
CC that the first individual has carcinoma. A method of inhibiting the
CC activity of a CAP comprises binding an inhibitor to the CAP. Treating
CC carcinomas comprises administering to a patient an inhibitor of CAP.
CC Neutralising the effect of a CAP comprises contacting an agent specific
CC for the CAP. The polypeptide specifically binds to the protein encoded by
CC the nucleic acid. It comprises an antibody that specifically binds to the
CC protein encoded by the nucleic acid. The nucleic acids are useful for
CC preparing a composition for diagnosing or treating carcinoma e.g.,
CC leukaemia or lymphoma. This sequence represents a human carcinoma
CC associated (CA) nucleic acid of the invention. Note: The sequence data
CC for this patent did not form part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 44075 BP; 8314 A; 11402 C; 11137 G; 8371 T; 0 U; 4851 Other;

Query Match 37.2%; Score 148.6; DB 12; Length 44075;

Best Local Similarity 74.9%; Pred. No. 7.7e-31;
Matches 200; Conservative 1; Mismatches 60; Indels 6; Gaps 1;

```
Qy        83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
             ||||||||||||||||||| |||||||||||| || ||||||||||||||||    | ||
Db      8665 GCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTAGGTGGATCACCTGAGGTCAGAGGTTC 8606

Qy       143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
             ||||||  |||||||||||||||||||||| || ||||||||||||||||| ||||||
Db      8605 AAGACCAGTCTGGCCAACATGGTGAAACCCCGCCTCTACTAAAAATACAAAAATTAGCTG 8546

Qy       203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
             ||||||||||||  | |||||| |||||||||||  |||| | |||   | ||
Db      8545 GGCATGGTGGCAGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCT 8486

Qy       263 AGATCGCAGAGTGAGCCGAAATCACAGATC------ACAGAGTGAGCAGAGTGAGACKCC 316
             || | |||   |     |    ||  || |      ||   |  | ||||||||||:|
Db      8485 TGAACCCAGGAGGCAGAGGTTGCAGTGAGCACTCCAACCTGGGCAACAGAGTGAGACTCT 8426

Qy       317 GTCTCAAAAACAACAACAAAAAACAAA 343
             |||||||||| || || |||||| |||
Db      8425 GTCTCAAAAAAAAAAAAAAAAAAAAA 8399
```

RESULT 46
ACN45134
ID ACN45134 standard; DNA; 49745 BP.
XX
AC ACN45134;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genomic sequence hCG31443.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 1930; 0pp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
SQ Sequence 49745 BP; 10151 A; 14745 C; 14496 G; 10151 T; 0 U; 202 Other;

```
  Query Match             37.2%;  Score 148.6;  DB 11;  Length 49745;
  Best Local Similarity   73.2%;  Pred. No. 8e-31;
  Matches  219;  Conservative    1;  Mismatches   70;  Indels    9;  Gaps    2;

Qy         51 CATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCC 110
              | ||  ||| |  | || ||      |     | ||||||||||||||||||| |||||||
Db      14658 CCTTAGCAGCCAGGCCAGGAGTGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCT 14717

Qy        111 AAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAAC 170
              ||||||||||||||||||||||  ||| |||||||| |||| |||| |||| |||||
Db      14718 GAGGTGGGCGGATCACCTGAGGTCGGGAGCTCGAGACCAGCCTGACCAATATGGAGAAAC 14777

Qy        171 CCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCA 230
              || |||| |||||||||||| |||| |||||||| | |||||||  |||||| |||||
Db      14778 CCTGTCTCTACTAAAAATAC-AAAATTAGCTGGGCGTTGTGGCACATGCCTGTAATCCCA 14836

Qy        231 GCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGA 290
              ||||| | |||| |||| ||| ||||||  |||||||    | |  |    | |  ||
Db      14837 GCTACCCGGGAGGCGGAGGTTGTGGTGAGCCCAGATCGCGCCATTACACTCTAGCCTAGG 14896

Qy        291 TCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACC 349
              ||||         ||| || ||:| |||||||||| |  |||||| ||  ||||||
Db      14897 CAACAG--------GAGGGAAACTCTGTCTCAAAAAAACAAACAAACAAACAAAAAAAC 14947


RESULT 47
AEE04996_2
Continuation (3 of 5) of AEE04996 from base 200001 (Cancer-associated gene SEQ ID NO:314
WP Sequence split into 5 fragments  LOCUS AEE04996 Accession Aee04996
WP     Fragment Name      Begin     End
WP     AEE04996_0             1     110000
WP     AEE04996_1        100001     210000
WP     AEE04996_2        200001     310000
WP     AEE04996_3        300001     410000
WP     AEE04996_4        400001     436348

  Query Match             37.2%;  Score 148.6;  DB 14;  Length 110000;
  Best Local Similarity   73.3%;  Pred. No. 1e-30;
  Matches  206;  Conservative    1;  Mismatches   65;  Indels    9;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||| ||||||||||||||||||||||||  ||
Db      14812 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCGGATCACCTGAGGTCAGAAGC 14871

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||| |||||||||||||||||||||  ||| |||||||||||||||  ||||
Db      14872 TGGAGACCAGCCTGGCCAACATGGTGAAACCCGATCTCTACTAAAAATACAAAAATTAGC 14931

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||| || |||||  |||||||| |||||||||||| |||| | ||| | | ||
Db      14932 TGGGCGTGATGGCATGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGTGGGAGAATTG 14991

Qy        261 TGAGATCGC---------AGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAG 311
                || | |         ||  | ||     | |          |   |  | |||||||||
Db      14992 CTTGAACCCAGGAGGCAGAGGTTTAGTGCCATTGCACTCCAGCTTGGGCAACAGAGTGAG 15051

Qy        312 ACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATA 352
              ||:|| ||| ||||| || || |||||  ||| ||| | ||
Db      15052 ACTCCATCTTAAAAAAAAAAAAAAAAAGAAAAGAAAACTTA 15092


RESULT 48
AAD54538
ID   AAD54538 standard; DNA; 133893 BP.
XX
AC   AAD54538;
XX
DT   26-JUN-2003  (first entry)
XX
DE   Human phosphatidylinositol biphosphate (PIB) DNA #1.
XX
KW   Human; p53 pathway; phosphatidylinositol biphosphate; PIB; therapeutic;
KW   cancer; apoptosis; cell proliferation; cytostatic; ds.
XX
OS   Homo sapiens.
XX
PN   WO200299125-A1.
XX
PD   12-DEC-2002.
XX
PF   03-JUN-2002; 2002WO-US017526.
XX
PR   05-JUN-2001; 2001US-0296076P.
PR   10-OCT-2001; 2001US-0328605P.
PR   15-FEB-2002; 2002US-0357253P.
XX
PA   (EXEL-) EXELIXIS INC.
XX
PI   Friedman L,  Plowman GD,  Belvin M,  Francis-Lang H,  Li D,  Funke RP;
XX
DR   WPI; 2003-167348/16.
XX
PT   Identifying candidate p53 pathway-modulating agents, useful as
PT   therapeutic targets for disorders associated with defective p53 function,
PT   comprises screening for agents that modulate phosphatidylinositol
PT   biphosphate activity.
XX
PS   Disclosure; Page 39-111; 139pp; English.
XX
CC   The present invention relates to a method of identifying candidate p53
CC   pathway-modulating agents. The method involves screening for agents that
CC   modulate the activity of phosphatidylinositol biphosphate (PIB). The
CC   methods are useful for identifying candidate p53 pathway-modulating
CC   agents used as therapeutic targets for disorders associated with
CC   defective p53 function. They are also useful for modulating p53 pathway
CC   in a mammalian cell or for diagnosing or treating a disease associated
CC   with defective p53 function, e.g. cancers such as breast cancer, colon
CC   cancer, lung cancer or ovarian cancer. Sequences of the invention are
CC   useful for identifying and testing agents that modulate PIB function. The
CC   animal models are useful for in vivo assays to test the activity of
CC   candidate p53-modulating agents or to assess the role of PIB in a p53
CC   pathway process such as apoptosis or cell proliferation. The present
CC   sequence is human PIB DNA
XX
SQ   Sequence 133893 BP; 33343 A; 33622 C; 34524 G; 32404 T; 0 U; 0 Other;

  Query Match             37.2%;  Score 148.6;  DB 9;  Length 133893;
  Best Local Similarity   70.8%;  Pred. No. 1.1e-30;
  Matches  196;  Conservative    1;  Mismatches   80;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| ||||||| |||  ||||||||||||||||  |||
Db     127385 ACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGGCGGATCACCTGAGGTCTGGAGT 127444

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | || |||| |||||||||||||||||| | ||| ||||||||||||||||  ||||
Db     127445 TTGAAACCAGCCTGGCCAACATGGTGAAACCTCATCTCTACTAAAAATACAAAAATTAGC 127504

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||  ||||||| |||||||||||||| ||||||||| |||| | |||   | ||
Db     127505 TGGGTGTGGTGGCGCACACCTGTAGTCGCAGCTACTCCGGAGGCTGAGGCAGGAGAATCA 127564

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                | |   |   | |  |      ||  | |  |||  | | |||  |||  :  |  ||
Db     127565 CTTAAACCTGGGAGGTGGAGGTTGTACCAAGCGGAGATCGTGGAGACAGAGCAGGCCCCT 127624

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
                | || |||||||  | || ||||||  || | |
Db     127625 GTCTCAAAAAAACAAAACAAAACAAAACCCCAAAAAA 127661


RESULT 49
ADH76849
ID   ADH76849 standard; DNA; 122557 BP.
XX
```

```
AC   ADH76849;
XX
DT   22-APR-2004  (first entry)
XX
DE   Melanin-concentrating hormone receptor 1 locus clone.
XX
KW   melanin-concentrating hormone receptor 1; MCHR1; SNP;
KW   single nucleotide polymorphism; anorectic; gene therapy; obesity; gene;
KW   ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(91268,C)
FT                   /*tag= a
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(91276,G)
FT                   /*tag= b
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(91497,A)
FT                   /*tag= c
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(94443,A)
FT                   /*tag= d
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(95538,G)
FT                   /*tag= e
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(95767,C)
FT                   /*tag= f
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(96227,A)
FT                   /*tag= g
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(96797,A)
FT                   /*tag= h
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(97004,G)
FT                   /*tag= i
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(97723,T)
FT                   /*tag= j
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(98848,T)
FT                   /*tag= k
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(99142,A)
FT                   /*tag= l
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(99450,A)
FT                   /*tag= m
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(99803,G)
FT                   /*tag= n
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(100220,A)
FT                   /*tag= o
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(101352,T)
FT                   /*tag= p
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(103156,C)
FT                   /*tag= q
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(103270,A)
FT                   /*tag= r
FT                   /standard_name= "Single nucleotide polymorphism"
XX
PN   WO2003104489-A2.
XX
PD   18-DEC-2003.
XX
PF   05-JUN-2003; 2003WO-EP005917.
XX
PR   05-JUN-2002; 2002EP-00012569.
XX
PA   (UYPH-) UNIV PHILIPPS MARBURG.
XX
PI   Platzer M,  Platzer C,  Gudermann T,  Hebebrand J,  Hinney A;
PI   Reichwald K;
XX
DR   WPI; 2004-062377/06.
XX
PT   New diagnostic composition, useful for diagnosing obesity related to the
PT   presence of a molecular variant of the MCHR1 gene or a susceptibility to
PT   the disorder.
XX
PS   Claim 1; Page; 76pp; English.
XX
CC   The invention relates to a novel diagnostic polynucleotide composition.
CC   The polynucleotide composition comprises: a sequence encoding a
CC   polypeptide with defined sequences given in the specification; a sequence
CC   capable of hybridizing to a melanin-concentrating hormone receptor 1
CC   (MCHR1) gene; a polynucleotide encoding an MCHR1 polypeptide; or a
CC   sequence comprising one or more of the nucleotide exchanges (SNP's) given
CC   in the specification and at least 8 bases of surrounding sequence of the
CC   MCHR1 gene. The composition has anorectic activity. The polynucleotide
CC   composition may be used in gene therapy to treat the disorders of the
CC   invention. The composition is useful for diagnosing obesity related to
CC   the presence of a molecular variant of the MCHR1 gene or a susceptibility
CC   to the disorder. The MCHR1 protein or polynucleotide is useful for
CC   preparing a medicament for treating or preventing obesity related to the
CC   presence of a molecular variant of the MCHR1 gene. This polynucleotide
CC   represents the melanin-concentrating hormone receptor 1 locus clone of
CC   the invention. This sequence is not shown in the specification. It has
CC   been taken from the GenBank accession number Z86090 provided in the
CC   specification.
XX
SQ   Sequence 122557 BP; 34723 A; 28259 C; 27523 G; 32052 T; 0 U; 0 Other;

  Query Match             37.1%;  Score 148.4;  DB 12;  Length 122557;
  Best Local Similarity   67.8%;  Pred. No. 1.2e-30;
  Matches  221;  Conservative    1;  Mismatches  102;  Indels    2;  Gaps    1;

Qy         24 GCCATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATG 83
              ||  |  |||  |||  |   ||  || | |       |  | || |     |    |
Db      67983 GCTGTCAGAAGGCAAGAAGAGATTTGAAAATAAGGCCCAGGCCAGGCGCGGTGGCTCACA 68042

Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
              |||||||||||||||||| | |||||| |||  |||||||  ||||||||||| ||| ||
Db      68043 CCTGTAATCCCAGCACTTTGAGAGGCCGAGGCCGGCGGATAACCTGAGGTCAGGAGTTCC 68102

Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
              ||||||  |||||||||||||||||||||||||| ||||||||||||||||  ||||  |
Db      68103 AGACCAGACTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCAG 68162

Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
              ||||||||||  ||||||| ||||||||||  |||| | |||  || ||
Db      68163 GCATGGTGGCAGGCACCTGTAATCCCAGCTACTTGGGAGGCTGAGGCTGGAGAATTGCTT 68222

Qy        264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGA--GTGAGCAGAGTGAGACKCCGTCTC 321
              || | | |   |    |    ||  |  | |||  |||   ||||||| ||:|| ||||
Db      68223 GAACCCGGGAGGCAGAGGTTGCAGTGGGCCGAGATCGTGCCAAGAGTGAAACTCCCTCTC 68282

Qy        322 AAAAACAACAACAAAAAACAAAAAAA 347
              |||||  ||  ||  | |||  || ||||
Db      68283 AAAAAAAAAAAAAGAAAGAAAGAAAA 68308


RESULT 50
AAK79514
ID   AAK79514 standard; DNA; 9469 BP.
XX
AC   AAK79514;
XX
```

DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34326.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.

```
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS   Disclosure; SEQ ID NO 34326; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting the
CC   nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention
XX
SQ   Sequence 9469 BP; 2940 A; 2203 C; 1941 G; 2385 T; 0 U; 0 Other;

  Query Match             37.0%;  Score 148;  DB 4;  Length 9469;
  Best Local Similarity   74.2%;  Pred. No. 7.1e-31;
  Matches  187;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;

Qy         46 TAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGG 105
              || |||       || | | || |  |   |    |||||||||||||||||||| |||
Db       1141 TATGACGGAAAATGGTCGGCCAGGCGAGGTGGCTCATGCCTGTAATCCCAGCACTTTGGG 1200

Qy        106 AGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGT 165
              |||||||||||| |||||||||||||||| ||| || |||||| ||||||||||||||||
Db       1201 AGGCCAAGGTGGGAGGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGT 1260

Qy        166 GAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAG 225
              |||||||||||| |||||||||||||||| |||| |||||||||||||| | ||||||
Db       1261 GAAACCCCGTCTCTACTAAAAATACAAAAATTAGCAGGGCATGGTGGCACGCGCCTGTAA 1320

Qy        226 TCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATC 285
              |||||||||||||||| | |||   | ||         || | | |   | | |    |
Db       1321 TCCCAGCTACTCAGGAGGCTGAGGCAGAAGAATCACTTGAACCCGGGAGGCGGAGGTTGC 1380

Qy        286 ACAGATCACAGA 297
              |  || |  |||
Db       1381 AGTGAGCCAAGA 1392


Search completed: July 17, 2006, 20:16:38
Job time : 542 secs
```

GenCore version 5.1.9

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 20:16:54 ; Search time 4678 Seconds
(without alignments)
4781.473 Million cell updates/sec

Title: SEQ1-47502C
Perfect score: 399.6
Sequence: 1 ccaggtactcagccatgtgc.....tattgcaggagaccaaaaag 400

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : EST:*
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_htc:*
7: gb_est2:*
8: gb_est7:*
9: gb_est8:*
10: gb_est9:*
11: gb_gss1:*
12: gb_gss2:*
13: gb_gss3:*
14: gb_gss4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	153	38.3	598	11	AQ416484	AQ416484 RPCI-11-1
2	151.4	37.9	457	11	AQ936397	AQ936397 RPCI-11-S
3	151.4	37.9	617	5	CK430465	CK430465 oj49g05.y
4	151.2	37.8	801	3	BU568469	BU568469 AGENCOURT
c 5	151	37.8	357	1	AA405549	AA405549 zw39f03.r
6	151	37.8	401	7	BE149224	BE149224 RC2-HT025
7	149.6	37.4	513	8	CN272188	CN272188 170006000
8	149.6	37.4	613	13	CZ455807	CZ455807 MCF738H04
c 9	148.6	37.2	542	9	DB313291	DB313291 DB313291
c 10	148.4	37.1	614	11	AQ541320	AQ541320 RPCI-11-3
c 11	148.4	37.1	847	9	DA572262	DA572262 DA572262
c 12	148	37.0	691	14	AG170656	AG170656 Pan trogl
13	148	37.0	749	2	BF965924	BF965924 602277287
14	147.6	36.9	538	11	AQ341973	AQ341973 RPCI11-11
c 15	147.4	36.9	426	1	AA648957	AA648957 ns30g05.s
16	147.2	36.8	897	10	DR759556	DR759556 HESC4_105
c 17	147	36.8	486	9	DA500385	DA500385 DA500385
c 18	147	36.8	540	9	DA180817	DA180817 DA180817
c 19	147	36.8	545	2	BG258140	BG258140 602379239
c 20	147	36.8	547	9	DA158439	DA158439 DA158439
c 21	147	36.8	569	9	DB176347	DB176347 DB176347
c 22	147	36.8	581	9	DA808834	DA808834 DA808834
c 23	147	36.8	1085	2	BM542252	BM542252 AGENCOURT
c 24	146.8	36.7	536	11	AQ381551	AQ381551 RPCI11-16
c 25	146.8	36.7	775	3	BQ432755	BQ432755 AGENCOURT
c 26	146.6	36.7	469	1	AI753536	AI753536 cr11h09.x
c 27	146.6	36.7	490	1	AI434037	AI434037 ti36b10.x
c 28	146.6	36.7	667	14	AG156377	AG156377 Pan trogl
c 29	146.4	36.6	406	7	AW902341	AW902341 QV3-NN102
c 30	146.4	36.6	524	1	AI302156	AI302156 qn58a05.x
31	146.4	36.6	1459	6	CR860050	CR860050 Pongo pyg
c 32	146.2	36.6	302	8	CN268905	CN268905 170005977
c 33	146.2	36.6	419	3	BM712012	BM712012 UI-E-DW1-
c 34	146.2	36.6	680	8	CN351916	CN351916 170005322
c 35	146.2	36.6	685	7	BE379282	BE379282 601237929
c 36	146.2	36.6	805	8	CN411940	CN411940 170005322
c 37	145.6	36.4	323	1	AA836548	AA836548 od36h04.s
c 38	145.4	36.4	416	3	BQ267427	BQ267427 ik07f12.x
39	145.4	36.4	464	11	AQ598684	AQ598684 HS_5336_B
40	145.4	36.4	518	3	BQ269776	BQ269776 ik30a05.x
c 41	145.4	36.4	551	9	DA175472	DA175472 DA175472
42	145.2	36.3	324	10	DW412368	DW412368 HHAGE0130
c 43	145.2	36.3	498	9	DB330188	DB330188 DB330188
c 44	145.2	36.3	515	9	DA714821	DA714821 DA714821
c 45	145.2	36.3	555	11	AQ314853	AQ314853 RPCI11-96
46	145.2	36.3	591	4	CB309712	CB309712 AGENCOURT
47	145.2	36.3	649	9	DA408951	DA408951 DA408951
48	145.2	36.3	849	13	CZ459725	CZ459725 MCF748c16
49	145	36.3	550	9	DA112846	DA112846 DA112846
c 50	145	36.3	555	9	DA198015	DA198015 DA198015
c 51	145	36.3	573	9	DB161173	DB161173 DB161173
c 52	145	36.3	771	11	AQ751578	AQ751578 HS_5576_B
c 53	144.8	36.2	515	10	R34561	R34561 yg58h06.r1
54	144.8	36.2	1813	6	AF305828	AF305828 Homo sapi
c 55	144.6	36.2	461	9	DA307259	DA307259 DA307259
c 56	144.6	36.2	463	9	DA073911	DA073911 DA073911
c 57	144.6	36.2	500	4	CB153412	CB153412 K-EST0210
c 58	144.6	36.2	519	9	DA364226	DA364226 DA364226
c 59	144.6	36.2	522	9	DA165286	DA165286 DA165286
c 60	144.6	36.2	552	9	DA340929	DA340929 DA340929
c 61	144.6	36.2	553	9	DA154522	DA154522 DA154522
c 62	144.6	36.2	554	9	DA172069	DA172069 DA172069
c 63	144.6	36.2	554	9	DA179393	DA179393 DA179393
c 64	144.6	36.2	555	9	DA156859	DA156859 DA156859
c 65	144.6	36.2	555	9	DA178653	DA178653 DA178653
c 66	144.6	36.2	556	9	DA627130	DA627130 DA627130
c 67	144.6	36.2	557	9	DA136912	DA136912 DA136912
c 68	144.6	36.2	558	9	DA306944	DA306944 DA306944
c 69	144.6	36.2	558	9	DA355632	DA355632 DA355632
c 70	144.6	36.2	560	9	DA286477	DA286477 DA286477
c 71	144.6	36.2	561	9	DA131278	DA131278 DA131278
c 72	144.6	36.2	561	9	DA166557	DA166557 DA166557
c 73	144.6	36.2	562	9	DA922190	DA922190 DA922190
c 74	144.6	36.2	563	9	DA158793	DA158793 DA158793
c 75	144.6	36.2	566	9	DA282948	DA282948 DA282948
c 76	144.6	36.2	568	9	DA254017	DA254017 DA254017
c 77	144.6	36.2	568	9	DA341427	DA341427 DA341427
c 78	144.6	36.2	571	9	DA165893	DA165893 DA165893
c 79	144.6	36.2	571	9	DA349302	DA349302 DA349302
c 80	144.6	36.2	572	9	DA175809	DA175809 DA175809
c 81	144.6	36.2	573	9	DA342756	DA342756 DA342756
c 82	144.6	36.2	573	9	DB168520	DB168520 DB168520
c 83	144.6	36.2	574	9	DB167344	DB167344 DB167344
c 84	144.6	36.2	575	9	DA304383	DA304383 DA304383
c 85	144.6	36.2	577	9	DA268711	DA268711 DA268711
c 86	144.6	36.2	577	9	DB170922	DB170922 DB170922
c 87	144.6	36.2	578	9	DA296828	DA296828 DA296828
c 88	144.6	36.2	578	9	DA350880	DA350880 DA350880
c 89	144.6	36.2	579	9	DA271020	DA271020 DA271020
c 90	144.6	36.2	584	9	DA253883	DA253883 DA253883
c 91	144.6	36.2	586	9	DA349834	DA349834 DA349834
c 92	144.6	36.2	1036	2	BM542055	BM542055 AGENCOURT

	93	144.4	36.1	301	3	BU957990 BU957990 AGENCOURT
	94	144.4	36.1	399	10	DW429644 DW429644 HHAGE0023
	95	144.4	36.1	505	9	DB278032 DB278032
c	96	144.4	36.1	735	9	DN994393 DN994393 TC111938
c	97	144.4	36.1	1004	14	DU799349 DU799349 fv01_fp00
c	98	144.4	36.1	1052	3	BM560673 BM560673 AGENCOURT
	99	144.2	36.1	500	14	AG028900 AG028900 Homo sapi
c	100	144.2	36.1	552	9	DB360773 DB360773
	101	144.2	36.1	692	14	AG146558 AG146558 Pan trogl
c	102	144.2	36.1	938	3	BU527958 BU527958 AGENCOURT
c	103	144	36.0	373	1	AA654778 AA654778 nt73g01.s
	104	144	36.0	716	13	CZ455149 CZ455149 MCF737J03
	105	144	36.0	895	2	BM452899 BM452899 AGENCOURT
c	106	143.8	36.0	647	14	AG048965 AG048965 Pan trogl
c	107	143.8	36.0	753	8	CO247422 CO247422 AGENCOURT
	108	143.6	35.9	324	10	DW412367 DW412367 HHAGE0130
c	109	143.6	35.9	511	9	DA873295 DA873295 DA873295
c	110	143.6	35.9	612	1	AA133332 zn92e02.s
	111	143.6	35.9	644	3	BQ183419 BQ183419 UI-H-EU0-
	112	143.6	35.9	662	14	AG037584 AG037584 Pan trogl
	113	143.6	35.9	673	3	BU736980 BU736980 UI-E-DW0-
c	114	143.6	35.9	673	11	AQ351427 AQ351427 RPCI11-11
	115	143.6	35.9	694	14	AG143386 AG143386 Pan trogl
	116	143.6	35.9	990	12	BZ601193 BZ601193 WHADA88TR
c	117	143.6	35.9	1035	2	BM543345 BM543345 AGENCOURT
	118	143.4	35.9	345	7	AW804959 AW804959 QV4-UM009
c	119	143.4	35.9	485	4	CA419549 CA419549 UI-H-FH0-
	120	143.4	35.9	693	14	AG179363 AG179363 Pan trogl
	121	143.4	35.9	718	8	CR789371 CR789371 DKFZp459J
c	122	143.4	35.9	831	2	BI758533 BI758533 603022857
	123	143.2	35.8	405	11	B64228 B64228 CIT-HSP-202
c	124	143.2	35.8	535	4	BX951790 BX951790 DKFZp781L
c	125	143.2	35.8	613	7	AW411283 AW411283 fh11d07.y
	126	143.2	35.8	897	14	DU796775 DU796775 fv01_fp00
	127	143.2	35.8	939	11	BH770745 BH770745 LLMGtag49
	128	143	35.8	289	11	B85944 B85944 RPCI11-21I1
	129	143	35.8	302	10	DW419716 DW419716 HHAGE0194
c	130	143	35.8	368	1	AA642809 AA642809 nu05a06.s
c	131	143	35.8	445	1	AA608602 AA608602 ae54f08.s
c	132	143	35.8	453	11	AQ020765 AQ020765 CIT-HSP-2
	133	143	35.8	508	11	AQ380981 AQ380981 RPCI11-16
	134	143	35.8	519	4	BY795043 BY795043 BY795043
c	135	143	35.8	562	9	DA343610 DA343610 DA343610
c	136	143	35.8	581	9	DA304206 DA304206 DA304206
	137	143	35.8	592	11	AQ318945 AQ318945 RPCI11-10
c	138	142.8	35.7	451	10	N27615 N27615 yw50a03.s1
c	139	142.8	35.7	589	9	DA128762 DA128762
c	140	142.8	35.7	610	4	CA442144 CA442144 UI-H-DI0-
	141	142.6	35.7	1945	6	BC035612 BC035612 Homo sapi
	142	142.4	35.6	374	11	AQ021361 AQ021361 CIT-HSP-2
	143	142.4	35.6	496	9	DB276584 DB276584 DB276584
c	144	142.4	35.6	585	3	BQ082129 BQ082129 K-EST0060
c	145	142.2	35.6	442	1	AA064978 AA064978 zf75f10.r
	146	142.2	35.6	450	11	AQ319210 AQ319210 RPCI11-98
	147	142.2	35.6	575	7	AV733228 AV733228 AV733228
c	148	142.2	35.6	761	9	DA834202 DA834202 DA834202
c	149	142	35.5	279	1	AI081241 AI081241 oy67b01.x
	150	142	35.5	360	1	AF236698 AF236698 AF236698

ALIGNMENTS

RESULT 1
AQ416484
LOCUS AQ416484 598 bp DNA linear GSS 23-MAR-1999
DEFINITION RPCI-11-153H2.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-153H2, genomic survey sequence.
ACCESSION AQ416484
VERSION AQ416484.1 GI:4470608
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 598)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: RPCI11-153H2.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: SP6
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .598
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7558537"
/db_xref="taxon:9606"
/clone="RPCI-11-153H2"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"
ORIGIN

Query Match 38.3%; Score 153; DB 11; Length 598;
Best Local Similarity 74.7%; Pred. No. 1.9e-19;
Matches 207; Conservative 1; Mismatches 61; Indels 8; Gaps 1;

```
Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
              ||||||||||||||||| |||||||||| ||||| |||||||||||||| ||| | |
Db        150 CCTGTAATCCCAGCACTTTGGGAGGCCAAGCTGGGCAGATCACCTGAGGTCAGGAGTTTG 209

Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
              | |||| ||||||||||||||||||||  |||| |||||||||||||||| |||||||
Db        210 ATACCAGCCTGGCCAACATGGTGAAACCTTGTCTCTACTAAAAATACAAAAATTAGCTGG 269

Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
              || ||||||   ||| ||| || |||||||||||| |||| | ||| |||||||||||| ||
Db        270 GCGTGGTGGTGCACGCCTATAATCCCAGCTACTCGGGAGGCAGAGGTTGCAGTGAGCAGA 329

Qy        264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
              |||  |          | |   |||      | | |  | ||||||| ||:| |||||||
Db        330 GATTG--------TGCCATTGCACGCTAGCCTGGGCAACAAGAGTGAAACTCTGTCTCAA 381

Qy        324 AAACAACAACAAAAAACAAAAAAACCATAAGACATTG 360
              |||  | |  |||||  | || |  ||  |   | ||
Db        382 AAAATAAATAAAAAAGAATAATACACAAGATCAAGTG 418
```

RESULT 2
AQ936397
LOCUS AQ936397 457 bp DNA linear GSS 15-MAR-2001
DEFINITION RPCI-11-Seg1-4-200N9-T7 RPCI-11 Human segment 1-4 genomic library Homo sapiens genomic clone RPCI-11-Seg1-4-200N9-T7 similar to Human Chr 14q24.3, genomic survey sequence.
ACCESSION AQ936397
VERSION AQ936397.1 GI:6654471
KEYWORDS GSS.

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 457)
  AUTHORS   Osoegawa,K., Mammoser,A.G., Wu,C., Frengen,E., Zeng,C.,
            Catanese,J.J. and de Jong,P.J.
  TITLE     A bacterial artificial chromosome library for sequencing the
            complete human genome
  JOURNAL   Genome Res. 11 (3), 483-496 (2001)
   PUBMED   11230172
COMMENT     Contact: de Jong, P.J.
            Children's Hospital Oakland Research Institute
            747 Fifty second Street, Oakland, CA 94609-1809, USA
            Tel: 510 450 7911
            Fax: 510 450 7924
            Email: pdejong@mail.cho.org
            BAC end sequences. For clone availability please contact Pieter de
            Jong (pdejong@.mail.cho.org). BACPAC Resources WWW site:
            www.chori.org/bacpac
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .457
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-Seg1-4-200N9-T7"
                     /sex="Male"
                     /cell_type="Lymphocyte"
                     /clone_lib="RPCI-11 Human segment 1-4 genomic library"
                     /note="Vector: pBACe3.6; BAC clones in E. coli DH10B"
ORIGIN

  Query Match             37.9%;  Score 151.4;  DB 11;  Length 457;
  Best Local Similarity   75.8%;  Pred. No. 4e-19;
  Matches  213;  Conservative    1;  Mismatches   56;  Indels   11;  Gaps    2;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               ||||||||||||||||||||| |||||||| |||||||| ||||||||||||||||||||
Db          47 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACCTGAGGTCAAGAGA 106

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               ||||||| ||||||||||||||||||||||||||||| |||||||||||||||| ||||
Db         107 TCGAGACTATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 166

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-------GCCGGAGATTGC 253
               ||||| |||||| |  || ||||||||||||||||||| |||       ||  |||| |
Db         167 TGGGCGTGGTGGTATGCATCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAAGAGAGTTG 226

Qy         254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGAT----CACAGAGTGAGCAGAGTG 309
               ||| |      |||||| | || |  | |  ||||    ||          || || |
Db         227 CTTGAACCTGAGAGGCAGAGGGTGCAGTGAACCAAGATACTCCAGCCTAGTGACACAGCG 286

Qy         310 AGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCA 350
               ||||:|||||||||||| ||  |   |||| |||| | | |
Db         287 AGACTCCGTCTCAAAAAAAAANAANNAAAANAAAACACCTA 327


RESULT 3
CK430465
LOCUS       CK430465                617 bp    mRNA    linear   EST 06-JAN-2004
DEFINITION  oj49g05.y1 Human lacrimal gland, unamplified: oj Homo sapiens cDNA
            clone oj49g05 5', mRNA sequence.
ACCESSION   CK430465
VERSION     CK430465.1  GI:40678584
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 617)
  AUTHORS   Ozyildirim,A.M., Wistow,G.J., Gao,J., Wang,J., Dickinson,D.P.,
            Frierson,H.P. Jr and Laurie,G.W.
  TITLE     The lacrimal gland transcriptome is an unusually rich source of
            rare and poorly characterized gene transcripts
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 46 (5), 1572-1580 (2005)
   PUBMED   15851553
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 49  row: g  column: 05
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .617
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="oj49g05"
                     /tissue_type="lacrimal gland"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human lacrimal gland, unamplified: oj"
                     /note="Organ: Eye; Vector: pCMVSport6; RNA was extracted
                     from 2 human lacrimal glands. A directionally cloned cDNA
                     library in the pCMVSPORT6 vector(Life Technologies) was
                     constructed at Bioserve Biotechnology (Laurel MD)
                     essentially following the protocols of the SuperScript
                     Plasmid System full details of which are contained in the
                     manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. EST analysis
                     was performed on the unamplified library at the NIH
                     Intramural Sequencing Center (NISC)."
ORIGIN

  Query Match             37.9%;  Score 151.4;  DB 5;  Length 617;
  Best Local Similarity   76.4%;  Pred. No. 3.7e-19;
  Matches  201;  Conservative    1;  Mismatches   52;  Indels    9;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||||||||||||| |  |||||||||||  |||| |||||| |||| ||| |||
Db          89 ACGCCTGTAATCCCAGCATTGTGGGAGGCCAAGTCGGGCAGATCACTTGAGTTCAGGAGT 148

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               |||||||| |||||||||||||||||||| |||| ||||||||||||||||| ||||
Db         149 TCGAGACCAGCCTGGCCAACATGGTGAAACCCAGTCTCTACTAAAAATACAAAAATTAGC 208

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||||||||||||  ||| || ||| |||||||||||||| | |||  |||||| |||
Db         209 CAGGCATGGTGGCACATGCCTATAATCCTAGCTACTCAGGAGGCAGAGGCTGCAGTAAGC 268

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               |||||       |    |  | |  || | |          |||||||||||:| ||||
Db         269 CAAGATCATGCCATTGCACTCCAGCCTAGGTGA---------CAGAGTGAGACTCTGTCT 319

Qy         321 CAAAAACAACAACAAAAAACAAA 343
               |||||| || || |||||| |||
Db         320 CAAAAAAAAAAAAAAAAAAAAAA 342


RESULT 4
BU568469
LOCUS       BU568469                801 bp    mRNA    linear   EST 16-SEP-2002
DEFINITION  AGENCOURT_10405312 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6615547
            5', mRNA sequence.
ACCESSION   BU568469
VERSION     BU568469.1  GI:22918758
KEYWORDS    EST.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 801)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2857  row: m  column: 19
            High quality sequence stop: 553.
FEATURES             Location/Qualifiers
     source          1. .801
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6615547"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_82"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).   Average insert size
                     1.35 kb (range 0.9-4.0 kb).  14/15 colonies contained
                     inserts by PCR. This library was enriched for full-length
                     clones and was constructed by Clontech  Laboratories (Palo
                     Alto, CA)."
ORIGIN

  Query Match             37.8%;  Score 151.2;  DB 3;  Length 801;
  Best Local Similarity   74.6%;  Pred. No. 3.9e-19;
  Matches  206;  Conservative    1;  Mismatches   59;  Indels   10;  Gaps    1;

Qy          84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
               ||||||||||||||||||| |||| ||| ||| |||||||||||||||| |||| ||| | |
Db          17 CCTGTAATCCCAGCACTTTGGGAAGCCGAGGAGGGCGGATCACCTGAAGTCAGGAGTTTG 76

Qy         144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
               |||||| | ||||||||||||||||||||| ||| ||||||||||||||||| |||| |
Db          77 AGACCAGCTTGGCCAACATGGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGCCTG 136

Qy         204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
               |||||||||||    |||||| |||||||||| ||||| | ||| |||||||||| ||
Db         137 GCATGGTGGCATGTGCCTGTAATCCCAGCTACCGAGGAGGCAGAGGCTGCAGTGAGCCGA 196

Qy         264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
               | |||||             ||  ||   ||    | | | |||| |||||:|  ||||||
Db         197 GCTCGCA----------CCATTGCACTCCAGCCTGGGTGAAGAGCGAGACTCTATCTCAA 246

Qy         324 AAACAACAACAAAAAACAAAAAAACCATAAGACATT 359
               ||| |  || ||||   ||| |||  | || | | |
Db         247 AAAAATAAAAAAAAGTAAAATAAAATAAAATAAAAT 282


RESULT 5
AA405549/c
LOCUS       AA405549                357 bp    mRNA    linear   EST 17-MAY-1997
DEFINITION  zw39f03.r1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
            IMAGE:772445 5' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   AA405549
VERSION     AA405549.1  GI:2063141
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 357)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
            Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 322.
FEATURES             Location/Qualifiers
     source          1. .357
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:772445"
                     /dev_stage="8-9 weeks"
                     /lab_host="DH10B"
                     /clone_lib="Soares_total_fetus_Nb2HF8_9w"
                     /note="Vector: pT7T3D-PacI; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was prepared from mRNA obtained from
                     pooled 8-9 week (total) fetus material with a Not I -
                     oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCTTAATTTTTTTTTTTTTTTT 3']. 
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

  Query Match             37.8%;  Score 151;  DB 1;  Length 357;
  Best Local Similarity   77.0%;  Pred. No. 5e-19;
  Matches  184;  Conservative    0;  Mismatches   55;  Indels    0;  Gaps    0;

Qy          32 AACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAAT 91
               || ||  | || || ||       |   |    ||| | ||| | |   | |||||||||
Db         287 AATCCGGAGATAAAGAAATTGAGGCATGAATTGGCTGGGCACGGTCGTCACGCCTGTAAT 228

Qy          92 CCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATC 151
               | |||||||| |||||||| |||  || |||||||||||||||| ||| || |||| | |
Db         227 CTCAGCACTTTGGGAGGCCGAGGCAGGTGGATCACCTGAGGTCAGGAGTTCAAGACTAGC 168

Qy         152 CTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTG 211
               |||||||||||||||||||||||||| ||||||||||||||||  || |||||||| ||
Db         167 CTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATGAGTTGGGCATGTTG 108

Qy         212 GCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCA 270
               |||  |||||||| |||| ||||||| |||| | ||| ||||||||||  |||| |||
Db         107 GCATGCACCTGTAATCCCGGCTACTCGGGAGGCTGAGGTTGCAGTGAGCCAAGATTGCA 49


RESULT 6
BE149224
LOCUS       BE149224                401 bp    mRNA    linear   EST 21-JUN-2000
DEFINITION  RC2-HT0252-120200-014-c06 HT0252 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE149224
VERSION     BE149224.1  GI:8611948
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 401)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC2-HT0252-120
            200-014-c06&t3=2000-02-12&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 16
            High quality sequence stop: 401.
FEATURES             Location/Qualifiers
     source          1. .401
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HT0252"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             37.8%;  Score 151;  DB 7;  Length 401;
  Best Local Similarity   78.3%;  Pred. No. 4.9e-19;
  Matches  206;  Conservative    0;  Mismatches   55;  Indels    2;  Gaps    2;

Qy          84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
               ||||||||||||||||| | |||||| ||||||| ||||||||||||| | ||| | |
Db          42 CCTGTAATCCCAGCACTTTGAGAGGCCGAGGTGGGTGGATCACCTGAGGTTAGGAGTTTG 101

Qy         144 AGACCATCCTGGCCAACATG-GTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
               |||||| |||||||||||| || |||||||||| ||||||||||||||||| ||||
Db         102 AGACCAGCCTGGCCAACATGTGTAAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCA 161

Qy         203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
               ||| ||||||||   |||||||| |||||||||||||||||| | ||| ||||||||||
Db         162 GGCGTGGTGGCGTGCACCTGTAATCCCAGCTACTCAGGAGGCAGAGGTTGCAGTGAGCCA 221

Qy         263 AGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCA 322
               ||||||||         | |   |||       | | |||| |||||  |||| | ||||||
Db         222 AGATCGCACTACTGCACTACTGCACTCCAGCCTGGGTGA-CAGAGCAAGACACTGTCTCA 280

Qy         323 AAAACAACAACAAAAAACAAAAA 345
               |||||||  | | |||| | | |
Db         281 AAAACAAACAAACAAAAAACAGA 303


RESULT 7
CN272188
LOCUS       CN272188                 513 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000600002901 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN272188
VERSION     CN272188.1  GI:47288602
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 513)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 513   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .513
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, DMSO-treated H9 cell
                     line"
                     /clone_lib="GRN_PREHEP"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from DMSO-treated hES cell line H9 (p22) maintained in
                     feeder-free conditions"
ORIGIN

  Query Match             37.4%;  Score 149.6;  DB 8;  Length 513;
  Best Local Similarity   71.4%;  Pred. No. 8.6e-19;
  Matches  197;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||  |||||| |||||||||||| | | ||||||  || ||||| |||
Db          51 ATGCCTGTAATCCTGGCACTTTGGGAGGCCAAGGCGTGTGGATCATTTGCGGTCAGGAGT 110

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               | ||||||| |||||||||||| |||||||||| || || ||| | |||||||| ||||
Db         111 TTGAGACCAGCCTGGCCAACATGATGAAACCCCGCCTCTATTAAGAGTACAAAAATTAGC 170

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||||||||     |||||||||||||||||||||||| ||||| ||||||||||
Db         171 TGGGCATGGTGGCGAGTGCCTGTAGTCCCAGCTACTCAGGAGGCGGAGGTTGCAGTGAGT 230

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               ||||  | |           |    | |    | | ||||||||||  |
Db         231 TGAGGTTGTGCCACTGCACTCCAGCTTCGGTGACAGAGTGAGATGCTGTCTCAAAAAAAA 290

Qy         321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGAC 356
               ||||  || |  || ||||||| |||  | |  ||
Db         291 AAAAACAAAAACAAACAAACAAACAAAAAAAACAAC 326


RESULT 8
CZ455807
LOCUS       CZ455807                 613 bp    DNA     linear   GSS 20-OCT-2005
DEFINITION  MCF738H04TF Human MCF7 breast cancer cell line library (MCF7_1)
            Homo sapiens genomic clone MCF7_38_H04, genomic survey sequence.
ACCESSION   CZ455807
VERSION     CZ455807.1  GI:77932102
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 613)
  AUTHORS   Volik,S.V., Raphael,B.J., Huang,G.-Q., Murnane,J., Brebner,J.H.,
            Bajsarowicz,K., Paris,P., Tao,Q., Kowbel,D., Lapuk,A.V., Kuo,W.-L.,
            Shagin,D.A., Shagina,I.A., Magrane,G., Gray,J.W., Jan,F.-C., de
            Jong,P., Pevzner,P. and Collins,C.
  TITLE     Decoding the genomic architecture and high throughput detection of
            fusion transcripts in breast cancer cell lines: implications for a
            tumor genome project
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .613
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_38_H04"
                     /sex="female"
                     /clone_lib="Human MCF7 breast cancer cell line library
                     (MCF7_1)"
                     /note="Vector: pECBAC1; Site_1: HindIII; This library was
                     constructed from MCF7 breast cancer cell line by Amplicon
                     Express (http://www.genomex.com) using their standard
                     procedure."
ORIGIN

  Query Match             37.4%;  Score 149.6;  DB 13;  Length 613;
  Best Local Similarity   75.6%;  Pred. No. 8.3e-19;
  Matches  201;  Conservative    1;  Mismatches   55;  Indels    9;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||| ||||||| |||||||| |||   || |||||||| ||||||| |||
Db         87 ACGCCTGTAATCCTAGCACTTTGGGAGGCCGAGGCAGGTGGATCACTTGAGGTCAGGAGT 146

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              ||||||||| ||| | |||||||||||||||||||| |||||||||||||||| |||
Db        147 TCGAGACCAGCCTAGTCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGT 206

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||| |||| ||||| || |||||||||||||| | | |||| |||| ||||||||||||
Db        207 TGGACATGCTGGCAGACGCCTGTAGTCCCAGCTCCCCTGGAGGTGGAGCTTGCAGTGAGC 266

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||||| |          |   | |   | | |          ||||||||||:| ||||
Db        267 CCAGATCACGCCACTGCACTCCACCCTGGGTGA---------CAGAGTGAGACTCTGTCT 317

Qy        321 CAAAAACAACAACAAAAAACAAAAAA 346
              |||||| || |  ||| || | ||||
Db        318 CAAAAAAAAAAGAAAAGAAAAGAAAA 343


RESULT 9
DB313291/c
LOCUS       DB313291                542 bp    mRNA    linear   EST 04-DEC-2005
DEFINITION  DB313291 CTONG2 Homo sapiens cDNA clone CTONG2000516 3', mRNA
            sequence.
ACCESSION   DB313291
VERSION     DB313291.1  GI:83195301
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 542)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
  PUBMED    16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .542
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CTONG2000516"
                     /tissue_type="tongue, tumor tissue"
                     /clone_lib="CTONG2"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match             37.2%;  Score 148.6;  DB 9;  Length 542;
  Best Local Similarity   79.5%;  Pred. No. 1.3e-18;
  Matches  175;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy         83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
              |||||||||||||||||| || |||||||||  ||| |||||| |||||||| ||| ||
Db        506 GCCTGTAATCCCAGCACTTTGGNAGGCCAAGGCTGGCAGATCACTTGAGGTCAGGAGTTC 447

Qy        143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
              ||||||||||||||||||||||||||||||| ||| ||||||||||||||||| ||||||
Db        446 AAGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGCTG 387

Qy        203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
              ||| ||||||||||  |||||||||||||||| |  |||| | |||  ||||||||||| |
Db        386 GGCGTGGTGGCACATGCCTGTAGTCCCAGCTATTTGGGAGGCTGAGGCTGCAGTGAGCCG 327

Qy        263 AGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAG 302
              |||| ||          | | |   | ||| ||||||
Db        326 AGATTGCGCCACTGCACTCCAGCCTGGGCGACACAGTGAG 287


RESULT 10
AQ541320/c
LOCUS       AQ541320                614 bp    DNA     linear   GSS 19-MAY-1999
DEFINITION  RPCI-11-343C15.TV RPCI-11 Homo sapiens genomic clone
            RPCI-11-343C15, genomic survey sequence.
ACCESSION   AQ541320
VERSION     AQ541320.1  GI:4871776
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 614)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-343C15.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .614
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7631390"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-343C15"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             37.1%;  Score 148.4;  DB 11;  Length 614;
  Best Local Similarity   72.1%;  Pred. No. 1.4e-18;
  Matches  209;  Conservative    1;  Mismatches   72;  Indels    8;  Gaps    1;

Qy          84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
               |||||||||||||||||| |||||||||||||||||||||||||||||||| ||  | |
Db         447 CCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAGGAATTTG 388

Qy         144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
               |||||| |||||||||||||||||||||| | | ||||||||||||||||| |||||||
Db         387 AGACCAGCCTGGCCAACATGGTGAAACCCCATTTCTACTAAAAATACAAAAATTAGCTGG 328

Qy         204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
               || ||||||      |||||| ||||||||||||  ||| | ||| ||| |||||||
Db         327 GCGTGGTGGTGGGTGCCTGTAATCCCAGCTACTCGAGAGGCAGAGGTTGTAGTGAGC--- 271

Qy         264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
                 |  ||   ||  | |                |    ||||||||||:|| ||||||
Db         270 -----CCAAGATTGCACCACTGCACTCCAGTCTGGGCGACAGAGTGAGACTCCATCTCAA 216

Qy         324 AAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTT 373
               ||| || || |||| | || |||    |  | ||| ||     | |  || |
Db         215 AAAAAAAAAGAAAAGAAAAGAAAGAAAACATACACTGAAATTTTAGGGGT 166


RESULT 11
DA572262/c
LOCUS       DA572262                 847 bp    mRNA    linear   EST 07-NOV-2005
DEFINITION  DA572262 HEMBB1 Homo sapiens cDNA clone HEMBB1000862 5', mRNA
            sequence.
ACCESSION   DA572262
VERSION     DA572262.1  GI:81139202
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 847)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
  PUBMED    16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute (HRI); cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .847
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HEMBB1000862"
                     /tissue_type="whole embryo, mainly body"
                     /dev_stage="embryo, 10 weeks"
                     /clone_lib="HEMBB1"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match             37.1%;  Score 148.4;  DB 9;  Length 847;
  Best Local Similarity   76.8%;  Pred. No. 1.3e-18;
  Matches  195;  Conservative    1;  Mismatches   52;  Indels    6;  Gaps    1;

Qy          84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
               |||||||||||||||||| |||||||| ||| ||| | |||||||||||||  ||| ||
Db         530 CCTGTAATCCCAGCACTTTGGGAGGCCCAGGCGGGGGCATCACCTGAGGTCCGGAGTTCA 471

Qy         144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
               |||||| |||||||||||||||||||||||| |||||||||||||||| |||||||
Db         470 AGACCAGTCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGG 411

Qy         204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
               |||| |||||||||| ||  ||||||||||||||  |||| | ||| ||||| ||||| ||
Db         410 GCATAGTGGCACACGCCCATAGTCCCAGCTACTTGGGAGGCTGAGGTTGCACTGAGCAGA 351

Qy         264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
               |||| |          |   ||     |  |    ||||||||||:|| || |||
Db         350 GATCACGCCACTGCTGCACTCCA------GCCTGGGTGACAGAGTGAGACTCCATCACAA 297

Qy         324 AAACAACAACAAAA 337
               ||| || || ||||
Db         296 AAAAAAAAAAAAA 283


RESULT 12
AG170656/c
LOCUS       AG170656                 691 bp    DNA     linear   GSS 09-JAN-2002
DEFINITION  Pan troglodytes DNA, clone: RP43-039L04.TJ, genomic survey
            sequence.
ACCESSION   AG170656
VERSION     AG170656.1  GI:16700334
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library RPCI-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 691)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                Sequencing: TJ
             LIBRARY
                Vector    : pBACe3.6
                R.Site 1  : EcoRI
                R.Site 2  : EcoRI.
FEATURES            Location/Qualifiers
     source         1. .691
                    /organism="Pan troglodytes"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9598"
                    /clone="RP43-039L04.TJ"
                    /sex="male"
                    /cell_type="lymphocytes"
                    /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
ORIGIN

  Query Match             37.0%;  Score 148;  DB 14;  Length 691;
  Best Local Similarity   71.1%;  Pred. No. 1.7e-18;
  Matches  212;  Conservative    1;  Mismatches   76;  Indels    9;  Gaps    1;

Qy         45 ATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGG 104
              ||||    ||    |   ||| || |       |     | ||||| ||||||||||||| ||
Db        405 ATAAAGTTTTTAAAATGCAGCCAGGTGTGGTGGGTCACACCTGTCATCCCAGCACTTTGG 346

Qy        105 GAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGG 164
              |||||  ||   || ||||| | |||| ||| ||| | |||| || ||| | ||||||||
Db        345 GAGGCTGAGACTGGTGGATCGCTTGAGCTCAGGAGTTTGAGATCAGCCTAGGCAACATGG 286

Qy        165 TGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTA 224
              |||||||| |||| ||| |||||||||||||  ||||||| ||||||||||  ||||||||
Db        285 TGAAACCCTGTCTCTACAAAAAATACAAAACTTAGCTGGACATGGTGGCATGCACCTGTA 226

Qy        225 GTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAAT 284
              |||||||||||||||||| | ||||||||| ||||  |||| |||   |    |    |
Db        225 GTCCCAGCTACTCAGGAGGCTGAGATTGCAGCGAGCCAAGATTGCACCATTGCACTCCAG 166

Qy        285 CACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAA 342
              |   |  ||         |||||||||||:||||||||||| || || ||||||  ||
Db        165 CCTGGGCCA---------CAGAGTGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAA 117


RESULT 13
BF965924
LOCUS       BF965924                749 bp    mRNA    linear   EST 22-JAN-2001
DEFINITION  602277287F2 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4365117 5',
            mRNA sequence.
ACCESSION   BF965924
VERSION     BF965924.1  GI:12333139
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 749)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10013  row: m  column: 22
            High quality sequence stop: 583.
FEATURES            Location/Qualifiers
     source         1. .749
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:4365117"
                    /tissue_type="osteosarcoma, cell line"
                    /lab_host="DH10B (phage-resistant)"
                    /clone_lib="NIH_MGC_86"
                    /note="Organ: bone; Vector: pCMV-SPORT6; Site_1: NotI;
                    Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                    Average insert size 1.533 kb. Library enriched for
                    full-length clones and constructed by Life Technologies.
                    Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             37.0%;  Score 148;  DB 2;  Length 749;
  Best Local Similarity   74.3%;  Pred. No. 1.6e-18;
  Matches  199;  Conservative    1;  Mismatches   66;  Indels    2;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| |||||||  |||||||| |||||||||||||| | |||
Db        137 ACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGCAGATCACCTGAGGTTAGGAGT 196

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||| |||||||||||||||||||||||| ||| ||||||||||||||||| |||| .
Db        197 TAGAGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 256

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
                ||| ||||||| |||| |||||||||||||||||| ||||  |  ||     | ||
Db        257 CAGGCGTGGTGGCGCACATCTGTAGTCCCAGCTACTCGGGAG--GCTGAAGCACGAGAAT 314

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              ||       |  ||   || ||  |  |||    |   |     |||||  ||||:| ||||
Db        315 TGCTTGAACCCAGGAGGCAGAGGTTGCAGTGAGCCTGGGCGACAGAGCAAGACTCTGTCT 374

Qy        321 CAAAAACAACAACAAAAAACAAAAAAAC 348
              |||||| || || | | || ||| ||||
Db        375 CAAAAAAAAAAAAACAGAAGAAAGAAAC 402


RESULT 14
AQ341973
LOCUS       AQ341973                538 bp    DNA     linear   GSS 06-MAY-1999
DEFINITION  RPCI11-111D18.TV RPCI-11 Homo sapiens genomic clone RPCI-11-111D18,
            genomic survey sequence.
ACCESSION   AQ341973
VERSION     AQ341973.1  GI:4166869
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 538)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
```

```
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .538
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="GDB:7542329"
                      /db_xref="taxon:9606"
                      /clone="RPCI-11-111D18"
                      /sex="Male"
                      /cell_type="Lymphocytes"
                      /clone_lib="RPCI-11"
                      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                      RPCI11 Human Male BAC Library"
ORIGIN

  Query Match          36.9%;  Score 147.6;  DB 11;  Length 538;
  Best Local Similarity   74.6%;  Pred. No. 2.1e-18;
  Matches  197;  Conservative    1;  Mismatches   65;  Indels    1;  Gaps    1;

Qy          84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
               |||||||||||||||||| ||||| ||||||| || ||||||| |||||||| ||| ||
Db         163 CCTGTAATCCCAGCACTTTGGGAGACCAAGGTAGGTGGATCACTTGAGGTCAGGAGTTCC 222

Qy         144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
               |||||| ||| |||||||||||||||||||||||| ||||||||||||||||| |||||||
Db         223 AGACCAGCCTCGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGG 282

Qy         204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
               |||||||||||| |||||||||||||||||||||||| | |||   | || || |
Db         283 GCATGGTGGCACGCACCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAG-GATCACT 341

Qy         264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
               | |||   |   | |         |   | |||||||||||:| |||||||
Db         342 TGAGCCTGAGATCGTGCCACTGCACTCCAGCTTGGGCAACAGAGTGAGACTCTGTCTCAA 401

Qy         324 AAACAACAACAAAAAACAAAAAAA 347
               | |  |  |  || || |  || |
Db         402 ATAATAATAATAATAATAGTAATA 425


RESULT 15
AA648957/c
LOCUS       AA648957               426 bp    mRNA    linear   EST 29-OCT-1997
DEFINITION  ns30g05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1185176 3'
            similar to contains Alu repetitive element;contains element PTR5
            repetitive element ;, mRNA sequence.
ACCESSION   AA648957
VERSION     AA648957.1  GI:2575386
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 426)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            unknown library type
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 368.
FEATURES              Location/Qualifiers
     source           1. .426
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1185176"
                      /tissue_type="germinal center B cell"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_GCB1"
                      /note="Vector: pT7T3D-PacI; Site_1: Not I; Site_2: Eco RI;
                      1st strand cDNA was prepared from human tonsillar cells
                      enriched for germinal center B cells by flow sorting
                      (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr.
                      David Allman (NCI) and Dr. Gerald Marti (CBER).  cDNA
                      synthesis was primed with a Not I - oligo(dT) primer
                      [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3'
                      ]. Double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. Library
                      went through one round of normalization, and was
                      constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN

  Query Match          36.9%;  Score 147.4;  DB 1;  Length 426;
  Best Local Similarity   71.1%;  Pred. No. 2.4e-18;
  Matches  224;  Conservative    1;  Mismatches   82;  Indels    8;  Gaps    2;

Qy          37 AAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAG 96
               | || |  | |  | |  |    | | |||| |     |  ||||||||||||||||
Db         399 ACATTTACAGAGAAAAGAGAAGTGGCTGGCAGGTGTGGTGGCTCATGCCTGTAATCCCAG 340

Qy          97 CACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGC 156
               ||||| |||||||||||| |||| | || |||||||||| ||| | ||||||| ||||||
Db         339 CACTTTGGGAGGCCAAGGCGGGCAGGTCGCCTGAGGTCAGGAGTTGGAGACCAGCCTGGC 280

Qy         157 CAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACA 216
               ||||||||  ||| || |||| ||||||||||||||||||| ||||||||  ||||||  ||
Db         279 CAACATGGCAAAAACCTGTCTCTACTAAAAATACAAAAATTAGCTGGGTGTGGTGGTGCA 220

Qy         217 CACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG-------TGAGCTGAGATCGC 269
                 |||||| |||||||| ||| |||| | ||| | | ||       ||| |  |||  ||
Db         219 TGCCTGTAATCCCAGCTTCTCGGGAGGCTGAGGTAGGAGAATCACTTGAACCCAGAAGGC 160

Qy         270 AGAGTGAGCCGAAATCACAGATCACAGAGTGAG-CAGAGTGAGACKCCGTCTCAAAAACA 328
               |||   || |  | |  ||||||    | ||| |||| ||||||:|| |||||||||  |
Db         159 GGAGGTTGCAGTGAGCCGAGATCAACCTGGGAGACAGAATGAGACTCCATCTCAAAAAAA 100

Qy         329 ACAACAAAAAACAAA 343
               | || |||||| |||
Db          99 AAAAAAAAAAAAAAA 85


RESULT 16
DR759556
LOCUS       DR759556               897 bp    mRNA    linear   EST 21-JUL-2005
DEFINITION  HESC4_105_B07.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7968807 5', mRNA sequence.
ACCESSION   DR759556
VERSION     DR759556.1  GI:71052256
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 897)
```

```
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Daniela S. Gerhard, Ph.D.
          Office of Cancer Genomics
          National Cancer Institute / NIH
          Bldg. 31 Rm10A07 Bethesda, MD 20892
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: BresaGen, Inc.
           cDNA Library Preparation: Express Genomics, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
          University of Georgia
           Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM 17081  row: c  column: 13
          Seq primer: JENREV (CAGGAAACAGCTATGACC)
          High quality sequence stop: 897.
FEATURES             Location/Qualifiers
     source          1. .897
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7968807"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             36.8%;  Score 147.2;  DB 10;  Length 897;
  Best Local Similarity   75.4%;  Pred. No. 2.2e-18;
  Matches  196;  Conservative    1;  Mismatches   59;  Indels    4;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | ||||||||||| ||||||| ||||||| |||  || ||||||||||||||||| |||
Db        593 ACGCCTGTAATCCTAGCACTTTGGGAGGCCGAGGCAGGTGGATCACCTGAGGTCAGGAGT 652

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| ||||||||||||||||||| || ||| |||||||||||| ||| ||||
Db        653 TCGAGACCAGCCTGGCCAACATGGTGAAACTCCATCTCTACTAAAAATACAGAAATTAGC 712

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||| |||||||  ||  |||||| |||||||||||  |||| | |||   | ||
Db        713 TGGGTATGGTGGTGCATGCCTGTAATCCCAGCTACTAGGGAGGCTGAGGCAGGAGAATCA 772

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                || | |||| | |  |    || || |    ||    ||||||||||:| ||||
Db        773 CTTGAACCCAGAAGGTGAAGGGCACAGTGACC----AGGCGCCAGAGTGAGACTCTGTCT 828

Qy        321 CAAAAACAACAACAAAAAAC 340
              |||||| || || |||||||
Db        829 CAAAAAGAAAAAGAAAAAAC 848


RESULT 17

DA500385/c
LOCUS       DA500385                486 bp    mRNA    linear   EST 07-NOV-2005
DEFINITION  DA500385 FCBBF3 Homo sapiens cDNA clone FCBBF3014320 5', mRNA
            sequence.
ACCESSION   DA500385
VERSION     DA500385.1  GI:81150116
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 486)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
  PUBMED    16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .486
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="FCBBF3014320"
                     /tissue_type="brain"
                     /dev_stage="fetal"
                     /clone_lib="FCBBF3"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match             36.8%;  Score 147;  DB 9;  Length 486;
  Best Local Similarity   71.9%;  Pred. No. 2.7e-18;
  Matches  192;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||| ||||||||||||||| |||||||  |||  || ||||||||||||||||| |||
Db        375 ACGCCTATAATCCCAGCACTTTGGGAGGCTGAGGCCGGTGGATCACCTGAGGTCAGGAGT 316

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || |||||| |||||||||||  |||||||||| ||| |||||||||||||||||| ||||
Db        315 TCAAGACCAGCCTGGCCAACACAGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 256

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||| |||  ||  |||||| |||||||||||  |||| | |||  |||||||||||
Db        255 TGGGCATGATGGTGCATGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGTTGCAGTGAGC 196

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                |||||||| |       |   | |   | | ||||||  ||         |
Db        195 CAAGATCGCACAACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCGTCTCCCAAAAAA 136

Qy        321 CAAAAACAACAACAAAAAACAAAAAAA 347
              ||| | |  || |||| | ||| | |
Db        135 TAAATAAATAAATAAAATAAAAATACA 109
```

```
RESULT 18
DA180817/c
LOCUS       DA180817                 540 bp    mRNA    linear   EST 01-NOV-2005
DEFINITION  DA180817 BRAMY2 Homo sapiens cDNA clone BRAMY2045648 5', mRNA
            sequence.
ACCESSION   DA180817
VERSION     DA180817.1  GI:78555451
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 540)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BRAMY2045648"
                     /tissue_type="amygdala"
                     /clone_lib="BRAMY2"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match             36.8%;  Score 147;  DB 9;  Length 540;
  Best Local Similarity   71.9%;  Pred. No. 2.7e-18;
  Matches  192;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||| |||||||||||||| |||||||  |||  || ||||||||||||||||| |||
Db         367 ACGCCTATAATCCCAGCACTTTGGGAGGCTGAGGCCGGTGGATCACCTGAGGTCAGGAGT 308

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || |||||| |||||||||| |||||||| ||| ||||||||||||||||| ||||
Db         307 TCAAGACCAGCCTGGCCAACACAGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 248

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||| |||  ||  |||||| |||||||||||  |||| | ||| ||||||||||||
Db         247 TGGGCATGATGGTGCATGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGTTGCAGTGAGC 188

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               ||||||| |      |  | |   | | ||||||  ||          |
Db         187 CAAGATCGCACAACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCGTCTCCCAAAAAA 128

Qy         321 CAAAAACAACAACAAAAAACAAAAAAA 347
               ||| | |  || |||| | ||| | |
Db         127 TAAATAAATAAATAAAATAAAAATACA 101
```

```
RESULT 19
BG258140/c
LOCUS       BG258140                 545 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  602379239F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4510256 5',
            mRNA sequence.
ACCESSION   BG258140
VERSION     BG258140.1  GI:12767956
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 545)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10391  row: m  column: 09
            High quality sequence stop: 542.
FEATURES             Location/Qualifiers
     source          1. .545
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4510256"
                     /tissue_type="embryonal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_92"
                     /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 2.5 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             36.8%;  Score 147;  DB 2;  Length 545;
  Best Local Similarity   71.9%;  Pred. No. 2.7e-18;
  Matches  192;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||| |||||||||||||| |||||||  |||  || ||||||||||||||||| |||
Db         445 ACGCCTATAATCCCAGCACTTTGGGAGGCTGAGGCCGGTGGATCACCTGAGGTCAGGAGT 386

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || |||||| |||||||||| |||||||| ||| ||||||||||||||||| ||||
Db         385 TCAAGACCAGCCTGGCCAACACAGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 326

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||| |||  ||  |||||| |||||||||||  |||| | ||| ||||||||||||
Db         325 TGGGCATGATGGTGCATGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGTTGCAGTGAGC 266

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               ||||||| |      |  | |   | | ||||||  ||          |
Db         265 CAAGATCGCACAACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCGTCTCCCAAAAAA 206

Qy         321 CAAAAACAACAACAAAAAACAAAAAAA 347
               ||| | |  || |||| | ||| | |
Db         205 TAAATAAATAAATAAAATAAAAATACA 179


RESULT 20
DA158439/c
```

```
LOCUS       DA158439                547 bp    mRNA    linear   EST 30-OCT-2005
DEFINITION  DA158439 BRAMY2 Homo sapiens cDNA clone BRAMY2017998 5', mRNA
            sequence.
ACCESSION   DA158439
VERSION     DA158439.1  GI:78296613
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .547
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BRAMY2017998"
                     /tissue_type="amygdala"
                     /clone_lib="BRAMY2"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match             36.8%;  Score 147;  DB 9;  Length 547;
  Best Local Similarity   71.9%;  Pred. No. 2.7e-18;
  Matches  192;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||| |||||||||||||| |||||||  |||  || ||||||||||||||||| |||
Db         372 ACGCCTATAATCCCAGCACTTTGGGAGGCTGAGGCCGGTGGATCACCTGAGGTCAGGAGT 313

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               || |||||| |||||||||||  |||||||||| ||| ||||||||||||||||| ||||
Db         312 TCAAGACCAGCCTGGCCAACACAGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGCC 253

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||| |||  ||  |||||| |||||||||||  |||| | ||| ||||||||||||
Db         252 TGGGCATGATGGTGCATGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGTTGCAGTGAGC 193

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               ||||||||  |       |   | |   | | ||||||  ||           |
Db         192 CAAGATCGCACAACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCGTCTCCCAAAAAA 133

Qy         321 CAAAAACAACAACAAAAAACAAAAAAA 347
               ||| | |  || |||| | ||| | |
Db         132 TAAATAAATAAATAAAATAAAAATACA 106


RESULT 21
```

```
DB176347/c
LOCUS       DB176347                569 bp    mRNA    linear   EST 10-DEC-2005
DEFINITION  DB176347 TKIDN2 Homo sapiens cDNA clone TKIDN2012877 5', mRNA
            sequence.
ACCESSION   DB176347
VERSION     DB176347.1  GI:83514420
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 569)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .569
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TKIDN2012877"
                     /tissue_type="kidney, tumor tissue"
                     /clone_lib="TKIDN2"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match             36.8%;  Score 147;  DB 9;  Length 569;
  Best Local Similarity   71.9%;  Pred. No. 2.7e-18;
  Matches  192;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||| |||||||||||||| |||||||  |||  || ||||||||||||||||| |||
Db         386 ACGCCTATAATCCCAGCACTTTGGGAGGCTGAGGCCGGTGGATCACCTGAGGTCAGGAGT 327

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               || |||||| |||||||||||  |||||||||| ||| ||||||||||||||||| ||||
Db         326 TCAAGACCAGCCTGGCCAACACAGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGCC 267

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||| |||  ||  |||||| |||||||||||  |||| | ||| ||||||||||||
Db         266 TGGGCATGATGGTGCATGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGTTGCAGTGAGC 207

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               ||||||||  |       |   | |   | | ||||||  ||           |
Db         206 CAAGATCGCACAACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCGTCTCCCAAAAAA 147

Qy         321 CAAAAACAACAACAAAAAACAAAAAAA 347
               ||| | |  || |||| | ||| | |
Db         146 TAAATAAATAAATAAAATAAAAATACA 120
```

```
RESULT 22
DA808834/c
LOCUS       DA808834                581 bp    mRNA    linear   EST 11-NOV-2005
DEFINITION  DA808834 OCBBF3 Homo sapiens cDNA clone OCBBF3024884 5', mRNA
            sequence.
ACCESSION   DA808834
VERSION     DA808834.1  GI:82089716
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="OCBBF3024884"
                     /tissue_type="brain"
                     /dev_stage="fetal"
                     /clone_lib="OCBBF3"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match             36.8%;  Score 147;  DB 9;  Length 581;
  Best Local Similarity   71.9%;  Pred. No. 2.7e-18;
  Matches  192;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||| ||||||||||||||| |||||||   |||  || ||||||||||||||| |||
Db         579 ACGCCTATAATCCCAGCACTTTGGGAGGCTGAGGCCGGTGGATCACCTGAGGTCAGGAGT 520

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || |||||| ||||||||||  |||||||||| ||| ||||||||||||||||| ||||
Db         519 TCAAGACCAGCCTGGCCAACACAGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 460

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||| |||  ||  |||||| |||||||||||  |||| | ||| ||||||||||
Db         459 TGGGCATGATGGTGCATGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGTTGCAGTGAGC 400

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               |||||||| |      |   | |  | | ||||||  ||          |
Db         399 CAAGATCGCACAACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCGTCTCCCAAAAAA 340

Qy         321 CAAAAACAACAACAAAAAACAAAAAAA 347
               ||| | |  || |||| | ||| | |
Db         339 TAAATAAATAAATAAAATAAAAATACA 313
```

```
RESULT 23
BM542252/c
LOCUS       BM542252                1085 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6436650 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539645
            5', mRNA sequence.
ACCESSION   BM542252
VERSION     BM542252.1  GI:18771599
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1085)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12234  row: h  column: 14
            High quality sequence stop: 418.
FEATURES             Location/Qualifiers
     source          1. .1085
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5539645"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             36.8%;  Score 147;  DB 2;  Length 1085;
  Best Local Similarity   71.9%;  Pred. No. 2.4e-18;
  Matches  192;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||| ||||||||||||||| |||||||   |||  || ||||||||||||||| |||
Db         329 ACGCCTATAATCCCAGCACTTTGGGAGGCTGAGGCCGGTGGATCACCTGAGGTCAGGAGT 270

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || |||||| ||||||||||  |||||||||| ||| ||||||||||||||||| ||||
Db         269 TCAAGACCAGCCTGGCCAACACAGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 210

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||| |||  ||  |||||| |||||||||||  |||| | ||| ||||||||||
Db         209 TGGGCATGATGGTGCATGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGTTGCAGTGAGC 150

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               |||||||| |      |   | |  | | ||||||  ||          |
Db         149 CAAGATCGCACAACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCGTCTCCCAAAAAA 90

Qy         321 CAAAAACAACAACAAAAAACAAAAAAA 347
               ||| | |  || |||| | ||| | |
Db          89 TAAATAAATAAATAAAATAAAAATACA 63
```

```
RESULT 24
AQ381551/c
LOCUS       AQ381551                536 bp    DNA     linear   GSS 21-MAY-1999
```

```
DEFINITION  RPCI11-165E2.TV RPCI-11 Homo sapiens genomic clone RPCI-11-165E2,
            genomic survey sequence.
ACCESSION   AQ381551
VERSION     AQ381551.1  GI:4352574
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 536)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI11-165E2.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7563073"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-165E2"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             36.7%;  Score 146.8;  DB 11;  Length 536;
  Best Local Similarity   71.7%;  Pred. No. 2.9e-18;
  Matches  208;  Conservative    1;  Mismatches   73;  Indels    8;  Gaps    1;

Qy          84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
               |||||||||||||||||| |||||||||||||||||||||||| |||||||| ||  | |
Db         441 CCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCGGATCACTTGAGGTCAGGAATTTG 382

Qy         144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
               |||||| |||||||||||||||||||||| | | |||||||||||||||||| |||||||
Db         381 AGACCAGCCTGGCCAACATGGTGAAACCCCATTTCTACTAAAAATACAAAAATTAGCTGG 322

Qy         204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
               || ||||||      |||||| |||||||||||  ||| | ||| ||| |||||||
Db         321 GCGTGGTGGTGGGTGCCTGTAATCCCAGCTACTCGAGAGGCAGAGGTTGTAGTGAGC--- 265

Qy         264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
                   |  ||   ||   | |           |    ||||||||||:|| ||||||
Db         264 -----CCAAGATTGCACCACTGCACTCCAGTCTGGGCGACAGAGTGAGACTCCATCTCAA 210

Qy         324 AAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTT 373
               ||| || || |||| | || |||   |  | ||| ||    | |  || |
Db         209 AAAAAAAAAGAAAAGAAAAGAAAGAAAACATACACTGAAATTTTAGGGGT 160


RESULT 25
BQ432755/c
```

```
LOCUS       BQ432755                775 bp    mRNA    linear   EST 24-MAY-2002
DEFINITION  AGENCOURT_7860001 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:6108747
            5', mRNA sequence.
ACCESSION   BQ432755
VERSION     BQ432755.1  GI:21171831
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 775)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2353  row: i  column: 04
            High quality sequence stop: 555.
FEATURES             Location/Qualifiers
     source          1. .775
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6108747"
                     /tissue_type="from acute myelogenous leukemia"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_55"
                     /note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI
                     (ggccattatggcc); Double-stranded cDNA was prepared from
                     cell line RNA.  5' and 3' adaptors were used in cloning as
                     follows: 5' adaptor sequence:  5'-CACGGCCATTATGGCC-3' and
                     3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).   Average insert size
                     1.65 kb (range 0.9-4.0 kb).  14/15 colonies contained
                     inserts by PCR. This library was enriched for full-length
                     clones and was constructed by Clontech  Laboratories (Palo
                     Alto, CA)."
ORIGIN

  Query Match             36.7%;  Score 146.8;  DB 3;  Length 775;
  Best Local Similarity   72.5%;  Pred. No. 2.7e-18;
  Matches  190;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||||||| |||||||| |||||||| |||  || ||||||||||||||| |||
Db         645 ACGCCTGTAATCTCAGCACTTTGGGAGGCCGAGGCAGGTGGATCACCTGAGGTCAGGAGT 586

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || |||| | |||||||||||||||||||||||||||||| ||||||||||||||||  ||
Db         585 TCAAGACTAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATGAGT 526

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||| |||||  |||||||| |||| ||||||| |||| | ||| ||||||||||||
Db         525 TGGGCATGTTGGCATGCACCTGTAATCCCGGCTACTCGGGAGGCTGAGGTTGCAGTGAGC 466

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                 |||| |||        |   | |   |   | | ||||||     |
Db         465 CAAGATTGCACCACTGCACCCCAGCCTGGGCAAAACAGTGAGACTTTTTCTCAAAAAAAA 406

Qy         321 CAAAAACAACAACAAAAAACAA 342
               ||||| || || || | | ||
Db         405 AAAAAATAATAATAATAGAAAA 384
```

```
RESULT 26
AI753536/c
LOCUS       AI753536                469 bp    mRNA    linear   EST 20-JUN-2002
DEFINITION  cr11h09.x1 Human bone marrow stromal cells Homo sapiens cDNA clone
            HBMSC:cr11h09 3', mRNA sequence.
ACCESSION   AI753536
VERSION     AI753536.1  GI:5131800
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 469)
  AUTHORS   Jia,L., Young,M.F., Powell,J., Yang,L., Ho,N.C., Hotchkiss,R.,
            Robey,P.G. and Francomano,C.A.
  TITLE     Gene expression profile of human bone marrow stromal cells:
            high-throughput expressed sequence tag sequencing analysis
  JOURNAL   Genomics 79 (1), 7-17 (2002)
   PUBMED   11827452
COMMENT     Contact: Libin Jia
            Medical Genetics Branch
            National Human Genome Research Institute
            10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
            Tel: 301-402-4877
            Fax: 301-496-7157
            Email: libin@helix.nih.gov
            DNA Sequencing and analyses by National Institutes of Health
            Intramural Sequencing Center (NISC).
            Plate: 11  row: h  column: 09
            Seq primer: -21M13 forward primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .469
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HBMSC_cr11h09"
                     /sex="mixed"
                     /tissue_type="bone marrow stroma"
                     /dev_stage="mixed"
                     /lab_host="XL1-Blue MRF'/SOLR"
                     /clone_lib="Human bone marrow stromal cells"
                     /note="Vector: pBluescript; Site_1: EcoRI; Site_2: XhoI;
                     mRNA made from human bone marrow stroma, cDNA made by
                     oligo-dT priming. Directionally cloned. Size-selected for
                     average insert size >0.5 kb. Library constructed by Dr.
                     Marian Young and Dr. Pamela Gehron Robey (NIDCR). Library
                     supplied by Dr. Libin Jia (NHGRI)"
ORIGIN

  Query Match           36.7%;  Score 146.6;  DB 1;  Length 469;
  Best Local Similarity   67.9%;  Pred. No. 3.3e-18;
  Matches  218;  Conservative    1;  Mismatches   94;  Indels    8;  Gaps    1;

Qy         33 ACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATC 92
              || ||| | |  || |  | |     ||  | ||||      |    | ||||||||||
Db        321 ACNCAATTTTACATTAAGAAATACTGTGCAGGCCATGNGTGGTGGNTCAGGCCTGTAATC 262

Qy         93 CCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCC 152
              ||||||||| |||||||| |||||||| ||||||| |||||||| ||| |||||||||  ||
Db        261 CCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACCGGAGGTCAGGAGTTCGAGACCAGCC 202

Qy        153 TGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGG 212
              | |||||||| |||||||| |||| |||||||||||||||||| |||| ||||||||||
Db        201 TTGCCAACATAGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCCGGGCATGGTGG 142

Qy        213 CACACACCTGTAGTCCCAGCTACTCAGGAG--------CCGGAGATTGCAGTGAGCTGAG 264
              ||  |||||||| ||||||||||  ||||      | ||   |||| |      ||
Db        141 CAGGCACCTGTAATCCCAGCTACTAGGGAGGCTTTTGAACCCAGGAGGCAGAGGTTGCAG 82

Qy        265 ATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAA 324
                || |||   ||   | |          |   |     ||||||||| :| ||||| ||
```

```
Db         81 CGAGCTGAGATCGCGCCACTGCACTCCAGCNTGGGTGATAGAGTGAGATTCAGTCTCCAA 22

Qy        325 AACAACAACAAAAAACAAAAA 345
              || || || |||||| |||||
Db         21 AAAAAAAAAAAAAAAAAAAAA 1


RESULT 27
AI434037/c
LOCUS       AI434037                490 bp    mRNA    linear   EST 30-MAR-1999
DEFINITION  ti36b10.x1 NCI_CGAP_Lym12 Homo sapiens cDNA clone IMAGE:2132539 3'
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AI434037
VERSION     AI434037.1  GI:4293374
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 490)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Life Technologies catalog #: 11547-015
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1365   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 297.
FEATURES             Location/Qualifiers
     source          1. .490
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2132539"
                     /tissue_type="lymphoma, follicular mixed small and large
                     cell"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lym12"
                     /note="Organ: lymph node; Vector: pCMV-SPORT6; Site_1:
                     SalI; Site_2: NotI; Cloned unidirectionally.  Primer:
                     Oligo dT.  Average insert size 1.25 kb. Life Technologies
                     catalog #: 11547-015"
ORIGIN

  Query Match           36.7%;  Score 146.6;  DB 1;  Length 490;
  Best Local Similarity   75.5%;  Pred. No. 3.3e-18;
  Matches  182;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy         63 AGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGA 122
              ||||  | |||  |    || ||||||||||||||||||| ||||||||||||||||||||
Db        285 AGGCCAGGCACGTGGCTCATCCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCGGA 226

Qy        123 TCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACT 182
              |||  |||||||||  |||  |||||||||  |||||||||||||||||||||||  |||  ||||
Db        225 TCATCTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTACT 166

Qy        183 AAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAG 242
              |||||||||||||  ||||||||   ||||  |||||  ||||||  ||||||||||||  ||||
Db        165 AAAAATACAAAAATTAGCTGGGTGTGGTAGCACATGCCTGTAATCCCAGCTACTCGGGAG 106

Qy        243 CCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAG 302
              | |||  | | ||         || | |||    |     |     ||  || |  |||  | |
Db        105 GCTGAGGTAGAAGAATCACTTGAACCCAGGAGGCAGAGGCTCCAGTGAGCCGAGATTACG 46

Qy        303 C 303
              |
```

```
Db         45 C 45


RESULT 28
AG156377/c
LOCUS       AG156377                667 bp    DNA     linear   GSS 09-JAN-2002
DEFINITION  Pan troglodytes DNA, clone: RP43-020P17.T7, genomic survey
            sequence.
ACCESSION   AG156377
VERSION     AG156377.1  GI:16686055
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library RPCI-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 667)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                 Sequencing: T7
             LIBRARY
                 Vector    : pBACe3.6
                 R.Site 1  : EcoRI
                 R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .667
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-020P17.T7"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
ORIGIN

  Query Match             36.7%;  Score 146.6;  DB 14;  Length 667;
  Best Local Similarity   78.2%;  Pred. No. 3.1e-18;
  Matches  176;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               ||||||||| |||||||||| |||||||  || ||||| ||||||||||||||||| |||
Db         559 ATGCCTGTAGTCCCAGCACTTTGGGAGGCTGAGATGGGCAGATCACCTGAGGTCAGGAGT 500

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               | ||||||| |||||||||||||||||||||| ||| ||||||||||||||||| ||||
Db         499 TTGAGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 440

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||||||||||||  | |||||||||||||||||||||||| | |||  | || |
Db         439 TGGGCATGGTGGCAGGCGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAGTTG 380

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAG 305
                 || | | |   | |  |    ||  || |  ||| ||||| |
Db         379 CTGGAACCCGGGAGGCGGAGTTTGCAGTGAGCTGAGATTGAGCCG 335


RESULT 29
AW902341/c
```

```
LOCUS       AW902341                406 bp    mRNA    linear   EST 24-MAY-2000
DEFINITION  QV3-NN1023-130500-179-b08 NN1023 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW902341
VERSION     AW902341.1  GI:8066546
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 406)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  PUBMED    10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=QV3-NN1023-130
            500-179-b08&t3=2000-05-13&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 22
            High quality sequence stop: 406.
FEATURES             Location/Qualifiers
     source          1. .406
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="NN1023"
                     /note="Organ: nervous_normal; Vector: pucl8; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             36.6%;  Score 146.4;  DB 7;  Length 406;
  Best Local Similarity   86.2%;  Pred. No. 3.7e-18;
  Matches  162;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||| ||||||||||||   |||||| |||  ||||||||||||||||||| |||
Db         301 ACGCCTGTGATCCCAGCACTTTCAGAGGCCGAGGCAGGCGGATCACCTGAGGTCAGGAGT 242

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               ||||||||| ||||||||| |||||||||||||||||| |||||||||||||||| ||||
Db         241 TCGAGACCAGCCTGGCCAATATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 182

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||| ||||||    |||||||||||||||||||||||||| | ||| |||| || |||
Db         181 TGGGCGTGGTGGTGGGCACCTGTAGTCCCAGCTACTCAGGAGGCAGAGGTTGCGGTAAGC 122

Qy         261 TGAGATCG 268
                ||||| |
Db         121 CGAGATAG 114
```

```
RESULT 30
AI302156/c
LOCUS       AI302156                524 bp    mRNA    linear   EST 03-DEC-1998
DEFINITION  qn58a05.x1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1902416 3'
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AI302156
VERSION     AI302156.1  GI:3961502
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 524)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 461.
FEATURES             Location/Qualifiers
     source          1. .524
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1902416"
                     /tissue_type="2 pooled tumors (clear cell type)"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid5"
                     /note="Organ: kIdney; Vector: pT7T3D-PacI; Site_1: Not I;
                     Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                     oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAATATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M. Fatima Bonaldo.  "
ORIGIN

  Query Match           36.6%;  Score 146.4;  DB 1;  Length 524;
  Best Local Similarity  72.6%;  Pred. No. 3.5e-18;
  Matches  209;  Conservative    1;  Mismatches   62;  Indels   16;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| ||||||| ||| |||||||||| |||| ||| |||
Db        450 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCAGTTGAGCTCAGGAGT 391

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||||||||||||||||||||||||| ||||||||||||||||| ||||
Db        390 TCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 331

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||||||| |||||||||||||||||||||| |||| | ||| | | || | |
Db        330 TGGGCATGGTGGTGCACACCTGTAGTCCCAGCTACTTGGGAGGCTGAGGTAGGAGAGTGG 271

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAG----------------CA 304
                || |   |   | | |   ||  || |  ||| | |                  ||
Db        270 CTTGAGCCTGGGAGGTGGAGGTTGCAGTGAGCCGAGACCGTGCCTCCAGCCTGGGTGACA 211

Qy        305 GAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATA 352
              ||| |||||:| |||||||||| || || || ||  | || ||  | |
Db        210 GAGCGAGACTCTGTCTCAAAAATAATAATAATAATAATAATAAAA 163
```

```
RESULT 31
CR860050
LOCUS       CR860050                1459 bp    mRNA    linear   HTC 12-NOV-2004
DEFINITION  Pongo pygmaeus mRNA; cDNA DKFZp470B1412 (from clone DKFZp470B1412).
ACCESSION   CR860050
VERSION     CR860050.1  GI:55730961
KEYWORDS    HTC.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 1459)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  CONSRTM   The German cDNA Consortium
  TITLE     Direct Submission
  JOURNAL   Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764
            Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by Agowa (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            This clone (DKFZp470B1412) is available at the RZPD Deutsches
            Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
            Please contact RZPD for ordering:
            http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp470B1412
            Further information about the clone and the sequencing project is
            available at http://mips.gsf.de/projects/cdna/.
FEATURES             Location/Qualifiers
     source          1. .1459
                     /organism="Pongo pygmaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9600"
                     /clone="DKFZp470B1412"
                     /tissue_type="liver"
                     /clone_lib="470 (synonym: pliv1). Vector pSport1_Sfi; host
                     DH10B; sites SfilA + SfilB"
                     /dev_stage="adult"
                     /note="TGF-beta receptor interacting protein 1 (Homo
                     sapiens)"
     gene            1. .1459
                     /gene="DKFZp470B1412"
     CDS             15. .992
                     /gene="DKFZp470B1412"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAH92198.1"
                     /db_xref="GI:55730962"
                     /translation="MKPILLQGHERSITQIKYNREGDLLFTVAKDPIVNVWYSVNGER
                     LGTYMGHTGAVWCVDADWDTKHVLTGSADNSCRLWDCETGKQLALLKTNSAVRTCGFD
                     FGGSIIMFSTDKQMGYQCFVSFFDLRDPSQIDNNEPYMKIPCNDSKITSAVWGPLGEC
                     IIAGHESGELNQYSAKSGEVLVNVKEHSRQINDIQLSRDMTMFVTASKDNTAKLFDST
                     TLEHQKTFRTERPVNSAALSPNYDHVVLGGGQEAMDVTTTSTRIGKFEARFFHLAFEE
                     EFGRVKGHFGLINSVAFHPDGKSYSSGGEDGYVRIHYFDPQYFEFEFEA"
ORIGIN

  Query Match           36.6%;  Score 146.4;  DB 6;  Length 1459;
  Best Local Similarity  73.8%;  Pred. No. 2.9e-18;
  Matches  186;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy         46 TAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGG 105
              |||||   ||     || | || |     |    ||||||||||||||| ||||| |||
Db        990 TAAGAAGCTGGATCTCCTGCCAGGCGCGGTGGCTCATGCCTGTAATCCCACCACTTTGGG 1049

Qy        106 AGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGT 165
              |||||||||||||||||||||||||||||||||| ||| | ||||||| |||| ||||||||||
Db       1050 AGGCCAAGGTGGGCGGATCACCTGAGGTCAGGAGTTTGAGACCAGCCTGACCAACATGGA 1109

Qy        166 GAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAG 225
              |||||||||||| |||||||||||||||||| |||| |||||||||||||||| ||| |||
```

```
Db       1110 GAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCGGGCATGGTGGCACACGCCTATAG 1169

Qy        226 TCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATC 285
              |||||||||||||||||| | |||   | ||         || | |||   |     |    |
Db       1170 TCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATCACTTGAACCCAGGAGGCAGAGGTTGC 1229

Qy        286 ACAGATCACAGA 297
              |  || |  |||
Db       1230 AGTGAGCTGAGA 1241


RESULT 32
CN268905/c
LOCUS       CN268905                 302 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000597786250 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN268905
VERSION     CN268905.1  GI:47285319
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 302)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 302   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .302
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, DMSO-treated H9 cell
                     line"
                     /clone_lib="GRN_PREHEP"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from DMSO-treated hES cell line H9 (p22) maintained in
                     feeder-free conditions"
ORIGIN

  Query Match             36.6%;  Score 146.2;  DB 8;  Length 302;
  Best Local Similarity   78.5%;  Pred. No. 4.3e-18;
  Matches  175;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| |||||||||||||| |||||||||||||||||||| |||
Db        299 ACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTAGGCGGATCACCTGAGGTCAGGAGT 240

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || |||||| |||||||||||||||||||||||||| |||||| |||||||||||| ||||
Db        239 TCAAGACCAGCCTGGCCAACATGGTGAAACCCCGTCCTTACTACAAATACAAAAATTAGC 180

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||| |||||||   | |||||||||||||||||||||||| | |||   | ||   |
Db        179 CGGGCGTGGTGGCGGGCGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATGG 120

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
              | || | | |   | |  |    ||  || |  ||| || ||
Db        119 CGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATTGCGC 77


RESULT 33
BM712012/c
LOCUS       BM712012                 419 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-DW1-ahc-a-08-0-UI.r1 UI-E-DW1 Homo sapiens cDNA clone
            UI-E-DW1-ahc-a-08-0-UI 5', mRNA sequence.
ACCESSION   BM712012
VERSION     BM712012.1  GI:19025270
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 419)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
             The following repetitive elements were found in this cDNA
            sequence: 1-181, >ALU (matched compliment) 55-316, >ALU (matched
            compliment)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .419
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-DW1-ahc-a-08-0-UI"
                     /tissue_type="lens"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-DW1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-DW1 is a normalized cDNA library containing the
                     following tissue(s): lens. The library was constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CGATTAGCGA. This library
                     was created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI)."
ORIGIN

  Query Match             36.6%;  Score 146.2;  DB 3;  Length 419;
  Best Local Similarity   75.3%;  Pred. No. 4e-18;
  Matches  195;  Conservative    1;  Mismatches   59;  Indels    4;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||| ||||||| |||||||| |||   || ||||||||||||||||| |||
Db        299 ACGCCTGTAATCCTAGCACTTTGGGAGGCCGAGGCAGGTGGATCACCTGAGGTCAGGAGT 240

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              |||||||| |||||||||||||||||||| || ||| |||||||||||||| ||| ||||
```

```
Db        239 TCGAGACCAGCCTGGCCAACATGGTGAAACTCCATCTCTACTAAAAATACAGAAATTAGC 180

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||| |||||||  ||  ||||||| ||||||||||  |||| | |||   | ||
Db        179 TGGGTATGGTGGTGCATGCCTGTAATCCCAGCTACTAGGGAGGCTGAGGCAGGAGAATCA 120

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                || | ||||  | |  |    ||  || |    ||    ||||||||||:| ||||
Db        119 CTTGAACCCAGAAGGTGAAGGGCACAGTGACC----AGGCGCCAGAGTGAGACTCTGTCT 64

Qy        321 CAAAAACAACAACAAAAAA 339
              |||||| || || ||||||
Db         63 CAAAAAGAAAAAGAAAAAA 45


RESULT 34
CN351916/c
LOCUS       CN351916                680 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000532267149 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN351916
VERSION     CN351916.1  GI:47351850
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 680   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, embryoid bodies
                     derived from H1, H7 and H9 cells"
                     /clone_lib="GRN_EB"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from embryoid body outgrowths derived from hES cell lines
                     H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
                     conditions."
ORIGIN

  Query Match            36.6%;  Score 146.2;  DB 8;  Length 680;
  Best Local Similarity  89.7%;  Pred. No. 3.7e-18;
  Matches  157;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||||| |||||| ||||||||||||||| |||
Db        437 ATGCCTGTAATCCCAGCACTTTGGGAGGCCAAAGTGGGCAGATCACCTGAGGTCAGGAGT 378

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| ||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        377 TCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 318

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG 255
              |||||||||||| ||||||||| |||||||||||| | || | |||   | ||
Db        317 TGGGCATGGTGGCGCACACCTGTAATCCCAGCTACTCCGAAGGCTGAGGCAGGAG 263


RESULT 35
BE379282/c
LOCUS       BE379282                685 bp    mRNA    linear   EST 21-JUL-2000
DEFINITION  601237929F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3610063 5',
            mRNA sequence.
ACCESSION   BE379282
VERSION     BE379282.1  GI:9324647
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM265  row: i  column: 08
            High quality sequence start: 4
            High quality sequence stop: 634.
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3610063"
                     /tissue_type="endometrium, adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_44"
                     /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California,  Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life  Technologies)."
ORIGIN

  Query Match            36.6%;  Score 146.2;  DB 7;  Length 685;
  Best Local Similarity  75.3%;  Pred. No. 3.7e-18;
  Matches  195;  Conservative    1;  Mismatches   59;  Indels    4;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||| ||||||| |||||||| |||   || ||||||||||||||| |||
Db        347 ACGCCTGTAATCCTAGCACTTTGGGAGGCCGAGGCAGGTGGATCACCTGAGGTCAGGAGT 288

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||||||||||||||||||| || |||| |||||||||||||| ||| ||||
Db        287 TCGAGACCAGCCTGGCCAACATGGTGAAACTCCATCTCTACTAAAAATACAGAAATTAGC 228

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||| |||||||  ||  |||||| ||||||||||  |||| | |||   | ||
Db        227 TGGGTATGGTGGTGCATGCCTGTAATCCCAGCTACTAGGGAGGCTGAGGCAGGAGAATCA 168

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                || | ||||  | |  |    ||  || |    ||    ||||||||||:| ||||
Db        167 CTTGAACCCAGAAGGTGAAGGGCACAGTGACC----AGGCGCCAGAGTGAGACTCTGTCT 112

Qy        321 CAAAAACAACAACAAAAAA 339
              |||||| || || ||||||
Db        111 CAAAAAGAAAAAGAAAAAA 93
```

```
RESULT 36
CN411940/c
LOCUS       CN411940                805 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000532275949 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN411940
VERSION     CN411940.1  GI:47399534
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 805)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
  PUBMED    15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 805   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .805
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, cell lines H1, H7, and
                     H9"
                     /clone_lib="GRN_ES"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from undifferentiated hES cell lines H1 (p32), H7 (p29),
                     and H9 (p26) maintained in feeder-free conditions"
ORIGIN

  Query Match             36.6%;  Score 146.2;  DB 8;  Length 805;
  Best Local Similarity   75.3%;  Pred. No. 3.6e-18;
  Matches  195;  Conservative    1;  Mismatches   59;  Indels    4;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||| ||||||| |||||||| |||  || ||||||||||||||| |||
Db         784 ACGCCTGTAATCCTAGCACTTTGGGAGGCCGAGGCAGGTGGATCACCTGAGGTCAGGAGT 725

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               ||||||||| |||||||||||||||||||| || ||| |||||||||||| ||| ||||
Db         724 TCGAGACCAGCCTGGCCAACATGGTGAAACTCCATCTCTACTAAAAATACAGAAATTAGC 665

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||| |||||||  ||  |||||| |||||||||||  |||| | |||   | ||
Db         664 TGGGTATGGTGGTGCATGCCTGTAATCCCAGCTACTAGGGAGGCTGAGGCAGGAGAATCA 605

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                  || | ||||  | |  |    ||  || |    ||    ||||||||||:| ||||
Db         604 CTTGAACCCAGAAGGTGAAGGGCACAGTGACC----AGGCGCCAGAGTGAGACTCTGTCT 549

Qy         321 CAAAAACAACAACAAAAAA 339
               |||||| || || ||||||
Db         548 CAAAAAGAAAAAGAAAAAA 530


RESULT 37
AA836548/c
LOCUS       AA836548                323 bp    mRNA    linear   EST 24-FEB-1998
DEFINITION  od36h04.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1370071
            similar to contains Alu repetitive element;contains element PTR5
            repetitive element ;, mRNA sequence.
ACCESSION   AA836548
VERSION     AA836548.1  GI:2910867
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 323)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 265.
FEATURES             Location/Qualifiers
     source          1. .323
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1370071"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_GCB1"
                     /note="Vector: pT7T3D-PacI; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was prepared from human tonsillar cells
                     enriched for germinal center B cells by flow sorting
                     (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr.
                     David Allman (NCI) and Dr. Gerald Marti (CBER).  cDNA
                     synthesis was primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN

  Query Match             36.4%;  Score 145.6;  DB 1;  Length 323;
  Best Local Similarity   76.0%;  Pred. No. 5.5e-18;
  Matches  206;  Conservative    1;  Mismatches   56;  Indels    8;  Gaps    2;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||| |||||||||||| || | | || |||||||||| |||
Db         284 ATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGNCAGGTCGCCTGAGGTCAGGAGT 225

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               | ||||||| |||||||||||||  ||| || |||| |||||||||||||||| ||||
Db         224 TGGAGACCAGCCTGGCCAACATGGCAAAAACCTGTCTCTACTAAAAATACAAAAATTAGC 165

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG----- 255
               ||||  ||||||  ||  |||||| |||||||| ||| |||| | ||| | | ||
Db         164 TGGGTGTGGTGGTGCATGCCTGTAATCCCAGCTTCTCGGGAGGCTGAGGTAGGAGAATCA 105

Qy         256 --TGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAG-CAGAGTGAGA 312
                 ||| |  |||  || |||    || |  | |  ||||||     | ||| |||| |||||
Db         104 CTTGAACCCAGAAGGCGGAGGTTGCAGTGAGCCGAGATCAACCTGGGAGACAGAATGAGA 45

Qy         313 CKCCGTCTCAAAAACAACAACAAAAAACAAA 343
               |:|| ||||||||| || || |||||| |||
Db          44 CTCCATCTCAAAAAAAAAAAAAAAAAAAAA 14
```

```
RESULT 38
BQ267427/c
LOCUS       BQ267427                416 bp    mRNA    linear   EST 15-JUL-2003
DEFINITION  ik07f12.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:5780494
            3', mRNA sequence.
ACCESSION   BQ267427
VERSION     BQ267427.1  GI:20492492
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 416)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40UP from Gibco
            High quality sequence stop: 355.
FEATURES             Location/Qualifiers
     source          1. .416
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5780494"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue  following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact  Hiroshi Inoue, MD/PhD for further
                     information on this library  (Metabolism Division, Permutt
                     Laboratory, Washington University School  of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110).  Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

  Query Match             36.4%;  Score 145.4;  DB 3;  Length 416;
  Best Local Similarity   78.0%;  Pred. No. 5.7e-18;
  Matches  202;  Conservative    1;  Mismatches   47;  Indels    9;  Gaps    2;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | ||||||||||| ||||||| |||||||||||||||||| |||||||||||||| |||
Db        250 ACGCCTGTAATCCTAGCACTTTGGGAGGCCAAGGTGGGCAGATCACCTGAGGTCAGGAGT 191

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||| |||| ||| |||||||||||||| || |||||||||||| ||||
Db        190 TCGAGACCAGCCTGACCAATATGATGAAACCCCGTCTCTA-AAAAAATACAAAAATTAGC 132

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||| ||||| ||  |||||||| |||||||||||||||| | ||| ||||||| ||
Db        131 CGGGCGTGGTGTCAGGCACCTGTAATCCCAGCTACTCAGGAGGCAGAGGTTGCAGTGGGC 72

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||| |||       | | |   |||      | | |  |  |||||||| |:| ||||
Db         71 TGAGGTCG--------GGCCATTGCACTCCAGCCTGGGCAAAAAGAGTGAGGCTCTGTCT 20

Qy        321 CAAAAACAACAACAAAAAA 339
              ||| || || || ||||||
Db         19 CAAGAAAAAAAAAAAAAAA 1


RESULT 39
AQ598684
LOCUS       AQ598684                464 bp    DNA     linear   GSS 08-JUN-1999
DEFINITION  HS_5336_B2_E09_SP6E RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=912 Col=18 Row=J, genomic survey sequence.
ACCESSION   AQ598684
VERSION     AQ598684.1  GI:5029896
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 464)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  PUBMED    10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 912  row: J  column: 18
            Seq primer: SP6
            Class: BAC ends
            High quality sequence stop: 464.
FEATURES             Location/Qualifiers
     source          1. .464
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=912 Col=18 Row=J"
                     /sex="male"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
ORIGIN

  Query Match             36.4%;  Score 145.4;  DB 11;  Length 464;
  Best Local Similarity   72.6%;  Pred. No. 5.6e-18;
  Matches  204;  Conservative    1;  Mismatches   67;  Indels    9;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||| |||||||||||||||||||||||||  ||
Db         97 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCGGATCACCTGAGGTCAGAAGC 156

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||| |||||||||||||||||||||  ||| |||||||||||||||| ||||
```

```
Db        157 TGGAGACCAGCCTGGCCAACATGGTGAAACCCGATCTCTACTAAAAATACAAAAATTAGC 216

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||| || |||||  |||||||| |||||||||||| |||| | ||| | | ||

Db        217 TGGGCGTGATGGCATGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGTGGGAGAATTG 276

Qy        261 TGAGATCGC---------AGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAG 311
                 || | |         ||  | ||    | |          |   |  | |||||||||

Db        277 CTTGAACCCAGGAGGCAGAGGTTTAGTGCCATTGCACTCCAGCTTGGGCAACAGAGTGAG 336

Qy        312 ACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATA 352
              ||:|| ||| | ||| || || |||||  ||| ||  | ||

Db        337 ACTCCATCTTATAAAAAAAAAAAAAAAGAAAAGAAGACTTA 377


RESULT 40
BQ269776
LOCUS       BQ269776                 518 bp    mRNA    linear   EST 15-JUL-2003
DEFINITION  ik30a05.x1 HR85 islet Homo sapiens cDNA clone IMAGE:5782185 3'
            similar to SW:ALU7_HUMAN P39194 ALU SUBFAMILY SQ SEQUENCE
            CONTAMINATION WARNING ENTRY. [1] ;, mRNA sequence.
ACCESSION   BQ269776
VERSION     BQ269776.1  GI:20494842
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 518)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40UP from Gibco
            High quality sequence stop: 446.
FEATURES             Location/Qualifiers
     source          1. .518
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5782185"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
ORIGIN
```

```
  Query Match           36.4%;  Score 145.4;  DB 3;  Length 518;
  Best Local Similarity   72.6%;  Pred. No. 5.5e-18;
  Matches  201;  Conservative    1;  Mismatches   72;  Indels    3;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| |||||||  |||    | | ||||||  | |||| |||
Db         62 ACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAAGTGCATCACCCAAAGTCACGAGT 121

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || |||||| |||||||||||||  |||| | |||| | |||||||||||||| ||||
Db        122 TCAAGACCAGCCTGGCCAACATGGCAAAACTCTGTCTCTGCTAAAAATACAAAAATTAGC 181

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
                 ||||| ||||||  |||||||| |||||||||||| |||| ||||| ||||||||||
Db        182 CAGGCATAGTGGCAGGCACCTGTAATCCCAGCTACTCTGGAGGCGGAGGTTGCAGTGAGC 241

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||||| |||         |   | |   | | ||||||     ||| |  | |:|
Db        242 TGAGATTGCACCACTGCTCTTTAGCCTGGGTGACAGAG---CAAGACTCTGTCTCAAAAA 298

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
              ||||| || || ||||  |||||| |  | || |||
Db        299 AAAAAAAAAAAAAAAAAGCCAAAAACAAAACAAAACA 335


RESULT 41
DA175472/c
LOCUS       DA175472                 551 bp    mRNA    linear   EST 02-NOV-2005
DEFINITION  DA175472 BRAMY2 Homo sapiens cDNA clone BRAMY2038827 5', mRNA
            sequence.
ACCESSION   DA175472
VERSION     DA175472.1  GI:78720791
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 551)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .551
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BRAMY2038827"
                     /tissue_type="amygdala"
                     /clone_lib="BRAMY2"
                     /note="Vector: pME18SFL3"
ORIGIN
```

```
  Query Match             36.4%;  Score 145.4;  DB 9;  Length 551;
  Best Local Similarity   71.5%;  Pred. No. 5.4e-18;
  Matches  191;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
Db        392 ACGCCTATAATCCCAGCACTTTGGGAGGCTGAGGCCGGTGGATCACCTGAGGTCAGGAGT 333

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
Db        332 TCAAGACCAGCCTGGCCAACACAGTGAACCCCCATCTCTACTAAAAATACAAAAATTAGC 273

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
Db        272 TGGGCATGATGGTGCATGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGTTGCAGTGAGC 213

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
Db        212 CAAGATCGCACAACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCGTCTCCCAAAAAA 153

Qy        321 CAAAAACAACAACAAAAAACAAAAAAA 347
Db        152 TAAATAAATAAATAAAATAAAAATACA 126


RESULT 42
DW412368
LOCUS       DW412368                 324 bp    mRNA    linear   EST 14-JAN-2006
DEFINITION  HHAGE013098 Human liver regeneration after partial hepatectomy Homo
            sapiens cDNA, mRNA sequence.
ACCESSION   DW412368
VERSION     DW412368.1  GI:84913924
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 324)
  AUTHORS   Xu,C.S.
  TITLE     Liver regeneration after PH
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Cun-Shuan Xu
            Henan Bioengineering Key Lab
            Henan Normal University
            No. 148 Jianshe Road, Xinxiang City, P.R.China
            Tel: 00863733328084
            Fax: 00863733326524
            Email: xucs@x263.net.
FEATURES             Location/Qualifiers
     source          1..324
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="liver"
                     /clone_lib="Human liver regeneration after partial
                     hepatectomy"
ORIGIN

  Query Match             36.3%;  Score 145.2;  DB 10;  Length 324;
  Best Local Similarity   71.1%;  Pred. No. 6.6e-18;
  Matches  209;  Conservative    0;  Mismatches   78;  Indels    7;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
Db         24 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCACTTGAGGTCAGGAGT 83

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
Db         84 TCGAGACCAGCCTGGTCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAACTAGC 143

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-------GCCGGAGATTGC 253
Db        144 CGAGCGTGGTGGTGCACGCTTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATTG 203

Qy        254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGAC 313
Db        204 CTTGAACCTGGGAGGCAGAGGTTGCAACCATTGCACTCCAGCCTGGGCGACAAGAGCAAG 263

Qy        314 KCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCT 367
Db        264 ACTCCTTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAACATGTCGGCCGCCT 317


RESULT 43
DB330188/c
LOCUS       DB330188                 498 bp    mRNA    linear   EST 10-DEC-2005
DEFINITION  DB330188 PROST2 Homo sapiens cDNA clone PROST2015815 3', mRNA
            sequence.
ACCESSION   DB330188
VERSION     DB330188.1  GI:83456206
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 498)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1..498
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PROST2015815"
                     /tissue_type="prostate"
                     /clone_lib="PROST2"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match             36.3%;  Score 145.2;  DB 9;  Length 498;
  Best Local Similarity   72.1%;  Pred. No. 6.1e-18;
  Matches  189;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy          2 CAGGTACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGGC 61
Db        341 CATTTTCTCATGCATCTATTGGTCATTTGTACATCATCTCTTGAGAAATGTCTTAAAAAA 282

Qy         62 CAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGG 121
Db        281 TGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGTGG 222
```

```
Qy        122 ATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTAC 181
              ||||||||||||||| ||| | |||||||  |||||||||||||||||||||||||| |||
Db        221 ATCACCTGAGGTCAGGAGTTTGAGACCAGTCTGGCCAACATGGTGAAACCCCGTCTCTAC 162

Qy        182 TAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA 241
              |||||||||||||| ||||||||| |||||||     |||||| ||||||||||||||||
Db        161 TAAAAATACAAAAATTAGCTGGGCGTGGTGGCGAGTGCCTGTAATCCCAGCTACTCAGGA 102

Qy        242 GCCGGAGATTGCAGTGAGCTGA 263
              | | |||| ||   | |  |||
Db        101 GTCTGAGACTGGGTTCATTTGA 80


RESULT 44
DA714821/c
LOCUS       DA714821                515 bp    mRNA    linear   EST 13-NOV-2005
DEFINITION  DA714821 NT2RI2 Homo sapiens cDNA clone NT2RI2021923 5', mRNA
            sequence.
ACCESSION   DA714821
VERSION     DA714821.1  GI:82370722
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 515)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
  PUBMED    16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .515
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="NT2RI2021923"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /clone_lib="NT2RI2"
                     /note="Vector: pME18SFL3; majorly NT2 neuron; mRNA from
                     NT2 neuronal precursor cells treated 2-weeks mitotic
                     inhibitor after 5-weeks retinoic acid (RA) induction."
ORIGIN

  Query Match             36.3%;  Score 145.2;  DB 9;  Length 515;
  Best Local Similarity   78.4%;  Pred. No. 6e-18;
  Matches  174;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||| |||||||||||||| |||||||  |||  || ||||||||||||||||| |||
Db        226 ACGCCTATAATCCCAGCACTTTGGGAGGCTGAGGCCGGTGGATCACCTGAGGTCAGGAGT 167

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              || |||||| ||||||||||  |||||||||| ||| ||||||||||||||||| ||||
Db        166 TCAAGACCAGCCTGGCCAACACAGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGCC 107

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||| |||  ||  |||||| |||||||||||  |||| | ||| ||||||||||
Db        106 TGGGCATGATGGTGCATGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGTTGCAGTGAGC 47

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAG 302
              ||||||||  |       |   | |   | | ||||||  ||
Db         46 CAAGATCGCACAACTGCACTCCAGCCTGGGTGACAGAGCAAG 5


RESULT 45
AQ314853/c
LOCUS       AQ314853                555 bp    DNA     linear   GSS 04-MAY-1999
DEFINITION  RPCI11-96D8.TV RPCI-11 Homo sapiens genomic clone RPCI-11-96D8,
            genomic survey sequence.
ACCESSION   AQ314853
VERSION     AQ314853.1  GI:4046316
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 555)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
  TITLE     Use of human BAC End Sequences for Sequence-Ready Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: RPCI11-96D8.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .555
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7536559"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-96D8"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             36.3%;  Score 145.2;  DB 11;  Length 555;
  Best Local Similarity   73.4%;  Pred. No. 5.9e-18;
  Matches  204;  Conservative    1;  Mismatches   59;  Indels   14;  Gaps    1;

Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
              |||||||||||||||||| |||||||| |||  |||||||||||||||||| ||| |||
Db        374 CCTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGGCGGATCACCTGAGGTCAGGAGTTCG 315

Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
              |||||| |||||||||||||||||||| || ||| |||||||||||||||||| |||||||
Db        314 AGACCAGCCTGGCCAACATGGTGAAACACCATCTCTACTAAAAATACAAAAATTAGCTGA 255
```

```
Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
              || |||||| | | |||||||||||||||||||||| |||| | |||    | ||
Db        254 GCGTGGTGACGCGCACCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTT 195

Qy        264 GATCGCAG--------------AGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTG 309
              || | |||               |||||||||| |||||     |          ||   ||||
Db        194 GAACCCAGGAGGCGGAGGTTGCAGTGAGCCGAGATCACGCCACTGCATTCCAGTCCAGTG 135

Qy        310 AGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAA 347
              ||||:|||||||||||| || || |  ||   ||  ||
Db        134 AGACTCCGTCTCAAAAAAAAAAAAAGTAAGTTAATTAA 97


RESULT 46
CB309712
LOCUS       CB309712                591 bp    mRNA    linear   EST 26-MAR-2004
DEFINITION  AGENCOURT_11828537 NICHD_Rh_Ov1 Macaca mulatta cDNA clone
            IMAGE:6913528 5', mRNA sequence.
ACCESSION   CB309712
VERSION     CB309712.1  GI:28832422
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Eliot Spindel
             cDNA Library Preparation: CLONTECH
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM3154  row: m  column: 15
            High quality sequence stop: 454.
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /db_xref="taxon:9544"
                     /clone="IMAGE:6913528"
                     /tissue_type="Ovary"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NICHD_Rh_Ov1"
                     /note="Organ: ovary; Vector: pDNR-LIB; Site_1: Sfi I;
                     Site_2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.0-4.0 kb. Tissue pooled from
                     pre-pubertal, post pubertal sn menopausal monkeys.
                     Constructed by Clontech. Note: this is a NICHD Library."
ORIGIN

  Query Match             36.3%;  Score 145.2;  DB 4;  Length 591;
  Best Local Similarity   84.8%;  Pred. No. 5.9e-18;
  Matches  162;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||||| |||||||||||||||||| |||||| ||||| || |||
Db        287 ATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCAGATCACTTGAGGACAGGAGT 346

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || |||||| |||||||||||||||| |||||| ||||||||||||||||||||||||||
Db        347 TCAAGACCAACCTGGCCAACATGGTGTAACCCCATCTTTACTAAAAATACAAAAAATAGC 406

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||||||||||  |||||||||| |||||||||||||| ||| | |||    | ||
Db        407 TGGGCATGGTGGCGGACACCTGTAATCCCAGCTACTCANGAGGCTGAGGCAGGAGAATCC 466

Qy        261 TGAGATCGCAG 271
              || | |||
Db        467 CTGGAACCCAG 477


RESULT 47
DA408951
LOCUS       DA408951                649 bp    mRNA    linear   EST 07-NOV-2005
DEFINITION  DA408951 BRTHA3 Homo sapiens cDNA clone BRTHA3015134 5', mRNA
            sequence.
ACCESSION   DA408951
VERSION     DA408951.1  GI:81121486
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 649)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
  PUBMED    16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .649
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BRTHA3015134"
                     /tissue_type="thalamus"
                     /clone_lib="BRTHA3"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match             36.3%;  Score 145.2;  DB 9;  Length 649;
  Best Local Similarity   70.1%;  Pred. No. 5.8e-18;
  Matches  195;  Conservative    0;  Mismatches   83;  Indels    0;  Gaps    0;

Qy         26 CATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCC 85
              | |||  | |    | |  |||||| || |  |    ||||  | |||  |    ||| |
Db        273 CCTGGACAACTCCCTCTCTATAAGAAATAGAAAAATGAGGCCAGGCACATGGCTCATGTC 332

Qy         86 TGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAG 145
              ||||||| |||||||||||||||  ||||| || |||||| |||||||| ||| |||||
Db        333 TGTAATCTCAGCACTTCGGGAGGCTGAGGTGAGCAGATCACTTGAGGTCAGGAGTTCGAG 392

Qy        146 ACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGC 205
              |||| || ||||||  |||  ||||||||||| ||||||||||||||||| |||||||||
Db        393 ACCAGCCCGGCCAATGTGGCAAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGC 452
```

```
Qy        206 ATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGA 265
              ||||| ||||||||||| ||||||||||   |||| | |||   | ||          ||
Db        453 GTGGTGACACACACCTGTAATCCCAGCTACTGGGGAGGCTGAGGCAGGAGAATTGCTTGA 512

Qy        266 TCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
              | ||||  | |  |    ||  || |  ||| |  ||
Db        513 GCCCAGAAGGCGGAGGTTGCAGTGAACCGAGATTATGC 550


RESULT 48
CZ459725
LOCUS       CZ459725                 849 bp    DNA     linear   GSS 20-OCT-2005
DEFINITION  MCF748c16TF Human MCF7 breast cancer cell line library (MCF7_1)
            Homo sapiens genomic clone MCF7_48c16, genomic survey sequence.
ACCESSION   CZ459725
VERSION     CZ459725.1  GI:77937953
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 849)
  AUTHORS   Volik,S.V., Raphael,B.J., Huang,G.-Q., Murnane,J., Brebner,J.H.,
            Bajsarowicz,K., Paris,P., Tao,Q., Kowbel,D., Lapuk,A.V., Kuo,W.-L.,
            Shagin,D.A., Shagina,I.A., Magrane,G., Gray,J.W., Jan,F.-C., de
            Jong,P., Pevzner,P. and Collins,C.
  TITLE     Decoding the genomic architecture and high throughput detection of
            fusion transcripts in breast cancer cell lines: implications for a
            tumor genome project
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .849
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_48c16"
                     /sex="female"
                     /clone_lib="Human MCF7 breast cancer cell line library
                     (MCF7_1)"
                     /note="Vector: pECBAC1; Site_1: HindIII; This library was
                     constructed from MCF7 breast cancer cell line by Amplicon
                     Express (http://www.genomex.com) using their standard
                     procedure."
ORIGIN

  Query Match            36.3%;  Score 145.2;  DB 13;  Length 849;
  Best Local Similarity  73.4%;  Pred. No. 5.5e-18;
  Matches  204;  Conservative    1;  Mismatches   59;  Indels   14;  Gaps    1;

Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
              |||||||||||||||||| |||||||| ||| ||||||||||||||||||| ||| |||
Db         70 CCTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGGCGGATCACCTGAGGTCAGGAGTTCG 129

Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
              |||||| |||||||||||||||||||| || ||| ||||||||||||||||| ||||||
Db        130 AGACCAGCCTGGCCAACATGGTGAAACACCATCTCTACTAAAAATACAAAAATTAGCTGA 189

Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
              || ||||| | | |||||||||||||||||||| |||| | |||   | ||
Db        190 GCGTGGTGACGCGCACCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTT 249

Qy        264 GATCGCAG--------------AGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTG 309
              || | |||               |||||||||| |||||    |         ||   ||||
Db        250 GAACCCAGGAGGCGGAGGTTGCAGTGAGCCGAGATCACGCCACTGCATTCCAGTCCAGTG 309

Qy        310 AGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
              ||||:|||||||||||| || || |  ||   ||  ||
Db        310 AGACTCCGTCTCAAAAAAAAAAAAAGTAAGTTAATTAA 347


RESULT 49
DA112846
LOCUS       DA112846                 550 bp    mRNA    linear   EST 03-NOV-2005
DEFINITION  DA112846 BRACE3 Homo sapiens cDNA clone BRACE3030866 5', mRNA
            sequence.
ACCESSION   DA112846
VERSION     DA112846.1  GI:79196234
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 550)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
  PUBMED    16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .550
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BRACE3030866"
                     /tissue_type="cerebellum"
                     /clone_lib="BRACE3"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match            36.3%;  Score 145;  DB 9;  Length 550;
  Best Local Similarity  76.5%;  Pred. No. 6.5e-18;
  Matches  218;  Conservative    1;  Mismatches   56;  Indels   10;  Gaps    3;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCA-CCTGAGGTCAAGAG 139
              | |||||||||||||||||| |||||||| ||||||| |||||| |||||||||| |||
Db        250 ACGCCTGTAATCCCAGCACTTAGGGAGGCCGAGGTGGGTGGATCACCCTGAGGTCAGGAG 309

Qy        140 ATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAAT-ACAAAAAATA 198
              | ||||||| |||||||||||||||||||||||||| ||| |||||||||| ||||||| ||
Db        310 TTTGAGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAAATCACAAAAATTA 369

Qy        199 GCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGA 258
              ||  ||| ||||| |||||||||| |||||||||||  ||||  |||| ||||||||||
Db        370 GCCAGGCGTGGTGATGCACACCTGTAATCCCAGCTACTCGGGAGGTGGAGGTTGCAGTGA 429
```

```
Qy        259 GCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGT 318
              || |||||||||         | |   |||       |   |  | ||||| |||||:| ||
Db        430 GCCGAGATCGCA--------CCACTGCACTTCCAGCCTGGGCAACAGAGCGAGACTCTGT 481

Qy        319 CTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCC 363
              ||||||||  | ||  ||||| ||| ||     ||  ||| ||||
Db        482 CTCAAAAAACAAAAACAAAAAAAAACAACGAGCAAACCATGGTCC 526


RESULT 50
DA198015/c
LOCUS       DA198015                 555 bp    mRNA    linear   EST 30-OCT-2005
DEFINITION  DA198015 BRAWH2 Homo sapiens cDNA clone BRAWH2000862 5', mRNA
            sequence.
ACCESSION   DA198015
VERSION     DA198015.1  GI:78309617
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 555)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
  TITLE     Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
  JOURNAL   Genome Res. 16 (1), 55-65 (2006)
   PUBMED   16344560
COMMENT     Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source          1. .555
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BRAWH2000862"
                     /tissue_type="brain"
                     /clone_lib="BRAWH2"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match             36.3%;  Score 145;  DB 9;  Length 555;
  Best Local Similarity   88.7%;  Pred. No. 6.5e-18;
  Matches  157;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||||| |||||| ||||||||||||||| |||
Db        487 ATGCCTGTAATCCCAGCACTTTGGGAGGCCAAAGTGGGCAGATCACCTGAGGTCAGGAGT 428

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| ||||||||||||||||||||||||| |||||||||||||||| ||||
Db        427 TCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 368

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTG 257
              |||||||||||| |||||||||| |||||||||||| | || | |||   | || |

Db        367 CGGGCATGGTGGCGCACACCTGTAATCCCAGCTACTCCGAAGGCTGAGGCAGGAGAG 311


Search completed: July 17, 2006, 22:20:11
Job time : 4693 secs
```

GenCore version 5.1.9

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 20:33:34 ; Search time 157 Seconds
(without alignments)
4767.156 Million cell updates/sec

Title: SEQ1-47502C
Perfect score: 399.6
Sequence: 1 ccaggtactcagccatgtgc.....tattgcaggagaccaaaaag 400

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : Issued_Patents_NA:*
1: /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS_COMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	Query % Match	Length	DB	ID	Description
c	1	159.6	39.9	601	3	US-09-949-016-192208	Sequence 192208,
c	2	158.4	39.6	10980	3	US-09-949-016-14471	Sequence 14471, A
c	3	158.4	39.6	15564	3	US-09-949-016-12783	Sequence 12783, A
	4	158.4	39.6	90776	3	US-09-949-016-17230	Sequence 17230, A
	5	154.4	38.6	44789	3	US-09-949-016-13909	Sequence 13909, A
c	6	154.4	38.6	173787	3	US-09-949-016-12542	Sequence 12542, A
c	7	154.4	38.6	173791	3	US-09-949-016-17302	Sequence 17302, A
c	8	153.6	38.4	71574	3	US-09-949-016-15580	Sequence 15580, A
c	9	153.4	38.4	152524	3	US-09-949-016-12683	Sequence 12683, A
c	10	153.4	38.4	152524	3	US-09-949-016-13194	Sequence 13194, A
c	11	152.8	38.2	40091	3	US-09-949-016-16011	Sequence 16011, A
c	12	152	38.0	601	3	US-09-949-016-69587	Sequence 69587, A
c	13	151.4	37.9	17050	3	US-09-949-016-12784	Sequence 12784, A
c	14	151.4	37.9	17050	3	US-09-949-016-13680	Sequence 13680, A
	15	150.8	37.7	20099	3	US-09-949-016-13784	Sequence 13784, A
c	16	150.4	37.6	19566	3	US-09-949-016-12096	Sequence 12096, A
c	17	150.4	37.6	19567	3	US-09-949-016-14114	Sequence 14114, A
	18	149.6	37.4	455726	3	US-09-949-016-14157	Sequence 14157, A
	19	149.6	37.4	481115	3	US-09-949-016-11940	Sequence 11940, A
c	20	149.4	37.4	157822	3	US-09-949-016-16723	Sequence 16723, A
	21	149	37.3	228896	3	US-09-949-016-17127	Sequence 17127, A
c	22	148.8	37.2	77994	3	US-09-949-016-12517	Sequence 12517, A
c	23	148.8	37.2	77994	3	US-09-949-016-16021	Sequence 16021, A
c	24	148.6	37.2	8100	3	US-09-949-016-13460	Sequence 13460, A
c	25	148.6	37.2	8100	3	US-09-949-016-13461	Sequence 13461, A
	26	148	37.0	601	3	US-09-949-016-152016	Sequence 152016,
c	27	147.8	37.0	601	3	US-09-949-002-2953	Sequence 2953, Ap
c	28	147.8	37.0	601	3	US-09-949-002-2954	Sequence 2954, Ap
c	29	147.8	37.0	601	3	US-09-949-002-4465	Sequence 4465, Ap
c	30	147.8	37.0	601	3	US-09-949-002-4466	Sequence 4466, Ap
	31	147.8	37.0	60595	3	US-09-949-002-650	Sequence 650, App
	32	147.8	37.0	60595	3	US-09-949-002-706	Sequence 706, App
	33	147.8	37.0	86945	3	US-09-949-016-13849	Sequence 13849, A
	34	147.8	37.0	86945	3	US-09-949-016-13850	Sequence 13850, A
	35	147.8	37.0	86945	3	US-09-949-016-13851	Sequence 13851, A
	36	147.8	37.0	86945	3	US-09-949-016-13852	Sequence 13852, A
	37	147.8	37.0	86945	3	US-09-949-016-13853	Sequence 13853, A
	38	147.8	37.0	86945	3	US-09-949-016-13854	Sequence 13854, A
	39	147.8	37.0	86945	3	US-09-949-016-13855	Sequence 13855, A
	40	147.8	37.0	86945	3	US-09-949-016-13856	Sequence 13856, A
	41	147.8	37.0	86945	3	US-09-949-016-13857	Sequence 13857, A
	42	147.8	37.0	86945	3	US-09-949-016-13858	Sequence 13858, A
	43	147.6	36.9	9798	3	US-09-949-016-14022	Sequence 14022, A
	44	147.6	36.9	9801	3	US-09-949-016-12819	Sequence 12819, A
	45	147.6	36.9	11820	5	US-09-936-271C-56	Sequence 56, Appl
c	46	147.4	36.9	26967	3	US-09-949-016-12926	Sequence 12926, A
c	47	147.4	36.9	35104	3	US-09-949-016-15831	Sequence 15831, A
	48	147.4	36.9	55114	3	US-09-949-016-16792	Sequence 16792, A
	49	147.4	36.9	116955	3	US-09-949-016-17565	Sequence 17565, A
c	50	147.2	36.8	301828	3	US-09-949-016-13969	Sequence 13969, A
c	51	147	36.8	601	3	US-09-949-016-41521	Sequence 41521, A
	52	146.6	36.7	161607	3	US-09-949-016-12210	Sequence 12210, A
c	53	146.4	36.6	228896	3	US-09-949-016-17127	Sequence 17127, A
c	54	146.2	36.6	50775	3	US-09-949-016-12858	Sequence 12858, A
c	55	146.2	36.6	50776	3	US-09-949-016-15438	Sequence 15438, A
c	56	146	36.5	122772	3	US-09-949-016-14132	Sequence 14132, A
c	57	145.8	36.5	27757	3	US-09-949-016-12407	Sequence 12407, A
c	58	145.8	36.5	27757	3	US-09-949-016-14545	Sequence 14545, A
c	59	145.8	36.5	101128	3	US-09-949-016-14293	Sequence 14293, A
c	60	145.6	36.4	66988	3	US-09-949-016-11942	Sequence 11942, A
c	61	145.6	36.4	66989	3	US-09-949-016-16063	Sequence 16063, A
	62	145.6	36.4	87774	3	US-09-949-016-12821	Sequence 12821, A
	63	145.4	36.4	21168	3	US-09-949-016-12643	Sequence 12643, A
	64	145.4	36.4	24546	3	US-09-949-016-16638	Sequence 16638, A
	65	145.4	36.4	77586	3	US-09-949-016-13220	Sequence 13220, A
	66	145.4	36.4	77586	3	US-09-949-016-13221	Sequence 13221, A
c	67	145.4	36.4	237241	3	US-09-949-016-16101	Sequence 16101, A
	68	145.2	36.3	601	3	US-09-949-016-152074	Sequence 152074,
	69	145.2	36.3	601	3	US-09-949-016-152147	Sequence 152147,
	70	145.2	36.3	601	3	US-09-949-016-152220	Sequence 152220,
	71	145.2	36.3	601	3	US-09-949-016-152293	Sequence 152293,
	72	145.2	36.3	601	3	US-09-949-016-159379	Sequence 159379,
	73	145.2	36.3	601	3	US-09-949-016-159452	Sequence 159452,
	74	145.2	36.3	601	3	US-09-949-016-159525	Sequence 159525,
	75	145.2	36.3	601	3	US-09-949-016-159598	Sequence 159598,
	76	145.2	36.3	32665	3	US-09-949-016-14098	Sequence 14098, A
	77	145.2	36.3	52667	3	US-09-949-016-12019	Sequence 12019, A
c	78	145.2	36.3	78846	3	US-09-949-016-12396	Sequence 12396, A
c	79	145.2	36.3	78846	3	US-09-949-016-12791	Sequence 12791, A
c	80	145.2	36.3	78846	3	US-09-949-016-12792	Sequence 12792, A
c	81	145.2	36.3	78846	3	US-09-949-016-12793	Sequence 12793, A
c	82	145.2	36.3	78850	3	US-09-949-016-16013	Sequence 16013, A
c	83	145.2	36.3	78850	3	US-09-949-016-16014	Sequence 16014, A
c	84	145.2	36.3	78850	3	US-09-949-016-16015	Sequence 16015, A
c	85	145.2	36.3	78850	3	US-09-949-016-16016	Sequence 16016, A
c	86	145.2	36.3	78850	3	US-09-949-016-16201	Sequence 16201, A
c	87	145.2	36.3	78850	3	US-09-949-016-16202	Sequence 16202, A
c	88	145.2	36.3	78850	3	US-09-949-016-16203	Sequence 16203, A
c	89	145.2	36.3	78850	3	US-09-949-016-16204	Sequence 16204, A
c	90	145.2	36.3	81001	3	US-09-750-580-1	Sequence 1, Appli
c	91	145.2	36.3	119032	3	US-09-949-016-12160	Sequence 12160, A
c	92	145.2	36.3	119032	3	US-09-949-016-17268	Sequence 17268, A
	93	145	36.3	601	3	US-09-949-016-155889	Sequence 155889,
c	94	145	36.3	41965	3	US-09-949-016-13067	Sequence 13067, A
c	95	145	36.3	87752	3	US-09-949-016-16807	Sequence 16807, A
c	96	144.8	36.2	601	3	US-09-949-016-196728	Sequence 196728,

c	97	144.8	36.2	19091	3	US-09-949-016-15805 Sequence 15805, A
c	98	144.8	36.2	43896	3	US-09-949-016-14942 Sequence 14942, A
c	99	144.8	36.2	84252	3	US-09-949-016-17315 Sequence 17315, A
	100	144.8	36.2	246230	3	US-09-949-016-17019 Sequence 17019, A
	101	144.8	36.2	246230	3	US-09-949-016-17020 Sequence 17020, A
	102	144.8	36.2	246230	3	US-09-949-016-17021 Sequence 17021, A
	103	144.8	36.2	246230	3	US-09-949-016-17022 Sequence 17022, A
c	104	144.6	36.2	601	3	US-09-949-016-195422 Sequence 195422,
	105	144.6	36.2	14961	3	US-09-949-016-13400 Sequence 13400, A
	106	144.6	36.2	36907	3	US-09-949-016-12633 Sequence 12633, A
	107	144.6	36.2	36913	3	US-09-949-016-15585 Sequence 15585, A
	108	144.6	36.2	43726	3	US-09-949-016-17578 Sequence 17578, A
c	109	144.6	36.2	47284	3	US-09-949-016-17029 Sequence 17029, A
c	110	144.6	36.2	57392	3	US-09-949-016-12070 Sequence 12070, A
c	111	144.6	36.2	57402	3	US-09-949-016-13293 Sequence 13293, A
c	112	144.6	36.2	61124	3	US-09-949-016-11914 Sequence 11914, A
c	113	144.6	36.2	61140	3	US-09-949-016-15771 Sequence 15771, A
c	114	144.6	36.2	131254	3	US-09-949-016-13734 Sequence 13734, A
c	115	144.6	36.2	157032	3	US-09-949-016-16502 Sequence 16502, A
c	116	144.6	36.2	346112	3	US-09-949-016-13165 Sequence 13165, A
	117	144.4	36.1	29717	3	US-09-949-016-16284 Sequence 16284, A
c	118	144.4	36.1	48039	3	US-09-949-016-15990 Sequence 15990, A
	119	144.4	36.1	60376	3	US-09-949-016-12423 Sequence 12423, A
	120	144.4	36.1	68702	3	US-09-949-016-16328 Sequence 16328, A
c	121	144.2	36.1	601	3	US-09-949-016-162811 Sequence 162811,
c	122	144.2	36.1	601	3	US-09-949-016-202637 Sequence 202637,
c	123	144.2	36.1	33155	3	US-09-949-016-16421 Sequence 16421, A
c	124	144.2	36.1	133358	3	US-09-949-016-16964 Sequence 16964, A
c	125	144.2	36.1	133360	3	US-09-949-016-12651 Sequence 12651, A
	126	144.2	36.1	199471	3	US-09-949-016-14083 Sequence 14083, A
c	127	144.2	36.1	312474	3	US-09-949-016-17434 Sequence 17434, A
	128	144	36.0	601	3	US-09-949-016-32649 Sequence 32649, A
	129	144	36.0	601	3	US-09-949-016-112168 Sequence 112168,
c	130	144	36.0	601	3	US-09-949-002-2952 Sequence 2952, Ap
c	131	144	36.0	601	3	US-09-949-002-4464 Sequence 4464, Ap
	132	144	36.0	15615	3	US-09-949-016-17221 Sequence 17221, A
	133	144	36.0	44166	3	US-09-949-016-15829 Sequence 15829, A
	134	144	36.0	54180	3	US-09-949-016-14894 Sequence 14894, A
	135	143.8	36.0	601	3	US-09-949-016-57694 Sequence 57694, A
	136	143.8	36.0	601	3	US-09-949-016-57695 Sequence 57695, A
	137	143.8	36.0	601	3	US-09-949-016-57696 Sequence 57696, A
c	138	143.8	36.0	7410	3	US-09-973-278-748 Sequence 748, App
c	139	143.8	36.0	7421	3	US-09-973-278-749 Sequence 749, App
	140	143.8	36.0	28283	3	US-09-949-016-15248 Sequence 15248, A
c	141	143.8	36.0	69874	3	US-09-949-016-12361 Sequence 12361, A
c	142	143.8	36.0	69874	3	US-09-949-016-13049 Sequence 13049, A
c	143	143.8	36.0	74804	3	US-09-949-016-15118 Sequence 15118, A
c	144	143.8	36.0	116955	3	US-09-949-016-17565 Sequence 17565, A
	145	143.8	36.0	236474	3	US-09-949-016-13418 Sequence 13418, A
	146	143.8	36.0	422592	3	US-09-949-016-14182 Sequence 14182, A
c	147	143.6	35.9	36952	3	US-09-949-016-14786 Sequence 14786, A
c	148	143.6	35.9	46794	3	US-09-949-016-12399 Sequence 12399, A
c	149	143.6	35.9	89584	3	US-09-949-016-17068 Sequence 17068, A
c	150	143.4	35.9	601	3	US-09-949-016-30346 Sequence 30346, A

ALIGNMENTS

```
RESULT 1
US-09-949-016-192208/c
; Sequence 192208, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 192208
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-192208

  Query Match             39.9%;  Score 159.6;  DB 3;  Length 601;
  Best Local Similarity   75.0%;  Pred. No. 1.2e-37;
  Matches  216;  Conservative    2;  Mismatches   56;  Indels   14;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||||| |||||||| ||| ||||||||  |||||||||| |||
Db        565 ACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATTTCCTGAGGTCAGGAGT 506

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || ||| || |||||||||||||||||||||||||||| ||||||||||||||||| ||||
Db        505 TCAAGATCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 446

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||||||||   |||||||||||||||||||| |||| | |||   | ||
Db        445 TGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCG 386

Qy        261 TGAGATCGCAG--------------AGTGAGCCGAAATCACAGATCACAGAGTGAGCAGA 306
              || | |||                |||||||||| ||| ||        | ||||  ||
Db        385 CTTGAACCCAGGCGGCGGAGGTTGCAGTGAGCCGAGATCGCACCAGCTTGGGTGACAAGC 326

Qy        307 GTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAG 354
              |||||||: | |||||||||| || :| |||| | |||| ||||| | |
Db        325 GTGAGACTTCATCTCAAAAAAAAAMAAAAAACAAAAAACAACCAGATG 278


RESULT 2
US-09-949-016-14471/c
; Sequence 14471, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14471
;   LENGTH: 10980
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14471

  Query Match             39.6%;  Score 158.4;  DB 3;  Length 10980;
  Best Local Similarity   73.9%;  Pred. No. 9e-37;
  Matches  201;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy         76 GCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCA 135
              | |  | ||||||||||||||||||  |||||||| ||| ||| ||||||||||| ||||
Db       4821 GGTTCACGCCTGTAATCCCAGCACTCTGGGAGGCCGAGGCGGGTGGATCACCTGATGTCA 4762

Qy        136 AGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAA 195
              ||| | |||| ||  ||||||||||||||||||||||||||| |||||||  |||||||||
Db       4761 GGAGTTTGAGATCAGTCTGGCCAACATGGTGAAACCCCGTCTCTACTAAATATACAAAAA 4702
```

```
Qy        196 ATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG 255
              ||||||| |||||| |||| | |||| |||||||||||| |||| ||||| ||||||
Db       4701 CCAGCTGGGTGTGGTGGTACACGCTTGTAATCCCAGCTACTCTGGAGGCGGAGGTTGCAG 4642

Qy        256 TGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKC 315
              ||||| || |  |||     |  |   | |   | | |||||| |||          |
Db       4641 TGAGCCGAAACTGCACCACTACACTCCAGCCTGGGTGACAGAGGGAGACTCTGTCTCCCA 4582

Qy        316 CGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
              |  :  ||||| || || ||||||| |||||||
Db       4581 GGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4550


RESULT 3
US-09-949-016-12783/c
; Sequence 12783, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12783
;   LENGTH: 15564
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12783

  Query Match             39.6%;  Score 158.4;  DB 3;  Length 15564;
  Best Local Similarity   73.9%;  Pred. No. 1e-36;
  Matches  201;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy         76 GCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCA 135
              | |  | ||||||||||||||||||  |||||||| ||| ||| |||||||||||| ||||
Db       9414 GGTTCACGCCTGTAATCCCAGCACTCTGGGAGGCCGAGGCGGGTGGATCACCTGATGTCA 9355

Qy        136 AGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAA 195
              ||| | |||| ||  ||||||||||||||||||||||||||||| |||||||| |||||||||
Db       9354 GGAGTTTGAGATCAGTCTGGCCAACATGGTGAAACCCCGTCTCTACTAAATATACAAAAA 9295

Qy        196 ATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG 255
              |||||||  |||||| |||| | |||| |||||||||||| |||| ||||| ||||||
Db       9294 CCAGCTGGGTGTGGTGGTACACGCTTGTAATCCCAGCTACTCTGGAGGCGGAGGTTGCAG 9235

Qy        256 TGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKC 315
              ||||| || |  |||     |  |   | |   | | |||||| |||          |
Db       9234 TGAGCCGAAACTGCACCACTACACTCCAGCCTGGGTGACAGAGGGAGACTCTGTCTCCCA 9175

Qy        316 CGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
              |    ||||| || || ||||||| |||||||
Db       9174 GGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 9143


RESULT 4
US-09-949-016-17230
; Sequence 17230, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17230
;   LENGTH: 90776
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17230

  Query Match             39.6%;  Score 158.4;  DB 3;  Length 90776;
  Best Local Similarity   75.0%;  Pred. No. 2.1e-36;
  Matches  216;  Conservative    1;  Mismatches   57;  Indels   14;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||||| |||||||| ||| ||||||||  |||||||||| |||
Db      10217 ACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATTTCCTGAGGTCAGGAGT 10276

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || ||| || |||||||||||||||||||||||||| |||||||||||||||| ||||
Db      10277 TCAAGATCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 10336

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||||||||||||   |||||||||||||||||||||| |||| | |||   | ||
Db      10337 TGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCG 10396

Qy        261 TGAGATCGCAG--------------AGTGAGCCGAAATCACAGATCACAGAGTGAGCAGA 306
                 || | |||              ||||||||| ||| ||      | ||||  ||
Db      10397 CTTGAACCCAGGCGGCGGAGGTTGCAGTGAGCCGAGATCGCACCAGCTTGGGTGACAAGC 10456

Qy        307 GTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAG 354
              |||||||: | |||||||||| ||   | |||| | |||| ||||| | |
Db      10457 GTGAGACTTCATCTCAAAAAAAAACAAAAAACAAAAAACAACCAGATG 10504


RESULT 5
US-09-949-016-13909
; Sequence 13909, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13909
;   LENGTH: 44789
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13909

  Query Match             38.6%;  Score 154.4;  DB 3;  Length 44789;
  Best Local Similarity   72.5%;  Pred. No. 2.4e-35;
  Matches  200;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              ||||||||||||  |||||| ||||||||||| | | ||||||  || ||||| |||
Db      14253 ATGCCTGTAATCCTGGCACTTTGGGAGGCCAAGGCGTGTGGATCATTTGCGGTCAGGAGT 14312
```

```
Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||| |||||||||||||| |||||||||| || || ||||||||||||||| ||||
Db      14313 TTGAGACCAGCCTGGCCAACATGATGAAACCCCGCCTCTATTAAAAATACAAAAATTAGC 14372

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||||||||     |||||||||||||||||||||||| ||||| ||||||||||
Db      14373 TGGGCATGGTGGCGAGTGCCTGTAGTCCCAGCTACTCAGGAGGCGGAGGTTGCAGTGAGT 14432

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||||| |          |   | |   | | ||||||||||  |
Db      14433 TGAGATTGTGCCACTGCACTCCAGCTTCGGTGACAGAGTGAGATGCTGTCTCAAAAAAAA 14492

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGAC 356
              ||||  || |  || ||||||| |||  | |  ||
Db      14493 AAAAACAAAAACAAACAAACAAACAAAAAAAACAAC 14528


RESULT 6
US-09-949-016-12542/c
; Sequence 12542, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12542
;   LENGTH: 173787
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12542

  Query Match            38.6%;  Score 154.4;  DB 3;  Length 173787;
  Best Local Similarity   76.9%;  Pred. No. 4.2e-35;
  Matches  203;  Conservative    1;  Mismatches   52;  Indels    8;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||| ||| |||||||||||||||| |||||||||||||||| |||
Db      18373 ACGCCTGTAATCCCAGCGCTTTGGGAGGCCAAGGTGGGAGGATCACCTGAGGTCAGGAGT 18314

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||| |||||||||||||||||||||  |||| ||| |||||||||||||| ||||
Db      18313 TGGAGACCAGCCTGGCCAACATGGTGAAACCTGGTCTCTACCAAAAATACAAAAATTAGC 18254

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||  |||||||   |||||||| |||||||||||||  ||| | |||  ||||||||||
Db      18253 TGGGCGTGGTGGCGGGCACCTGTAATCCCAGCTACTCCAGAGGCTGAGGCTGCAGTGAGC 18194

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||||||         |  | |   |||      | | | ||  ||||  | ||:|| |||
Db      18193 TGAGATC--------ATGCCACTGCACTCCAGCCTGGGCGACAAGAGCAAAACTCCATCT 18142

Qy        321 CAAAAACAACAACAAAAAACAAAA 344
              ||| || || || ||||||   ||
Db      18141 CAAGAAAAAAAAAAAAAAAGCCAA 18118


RESULT 7
US-09-949-016-17302/c
; Sequence 17302, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17302
;   LENGTH: 173791
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17302

  Query Match            38.6%;  Score 154.4;  DB 3;  Length 173791;
  Best Local Similarity   76.9%;  Pred. No. 4.2e-35;
  Matches  203;  Conservative    1;  Mismatches   52;  Indels    8;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||| ||| |||||||||||||||| |||||||||||||||| |||
Db      18373 ACGCCTGTAATCCCAGCGCTTTGGGAGGCCAAGGTGGGAGGATCACCTGAGGTCAGGAGT 18314

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||| |||||||||||||||||||||  |||| ||| |||||||||||||| ||||
Db      18313 TGGAGACCAGCCTGGCCAACATGGTGAAACCTGGTCTCTACCAAAAATACAAAAATTAGC 18254

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||  |||||||   |||||||| |||||||||||||  ||| | |||  ||||||||||
Db      18253 TGGGCGTGGTGGCGGGCACCTGTAATCCCAGCTACTCCAGAGGCTGAGGCTGCAGTGAGC 18194

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||||||         |  | |   |||      | | | ||  ||||  | ||:|| |||
Db      18193 TGAGATC--------ATGCCACTGCACTCCAGCCTGGGCGACAAGAGCAAAACTCCATCT 18142

Qy        321 CAAAAACAACAACAAAAAACAAAA 344
              ||| || || || ||||||   ||
Db      18141 CAAGAAAAAAAAAAAAAAAGCCAA 18118


RESULT 8
US-09-949-016-15580/c
; Sequence 15580, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15580
;   LENGTH: 71574
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(71574)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15580
```

```
  Query Match             38.4%;  Score 153.6;  DB 3;  Length 71574;
  Best Local Similarity   72.8%;  Pred. No. 5.1e-35;
  Matches  198;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | ||||||||| | | ||||| |||||||  | ||||||||||||||||||||||| |||
Db      57890 ACGCCTGTAATTCTACCACTTTGGGAGGCTGAAGTGGGCGGATCACCTGAGGTCAGGAGT 57831

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              ||||||||| ||||||||||||| |||| || |||| |||||||||||||||| ||||
Db      57830 TCGAGACCAGCCTGGCCAACATGGCGAAATCCTGTCTCTACTAAAAATACAAAAATTAGC 57771

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||  |||| |    | |||||| |||||||||||||||| ||||| ||||||||||
Db      57770 TGGGTGTGGTTGTGGGCGCCTGTAATCCCAGCTACTCAGGAGGCGGAGGTTGCAGTGAGC 57711

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              ||||||  |          |   | |   | | ||||||  ||
Db      57710 TGAGATTGTGCCACTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCAAAAAAAA 57651

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATA 352
              |||||  || ||||||||||||||  | | |
Db      57650 AAAAAAAAAAAACAAAAAACAAAAACCCAAAA 57619


RESULT 9
US-09-949-016-12683/c
; Sequence 12683, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12683
;   LENGTH: 152524
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12683

  Query Match             38.4%;  Score 153.4;  DB 3;  Length 152524;
  Best Local Similarity   69.8%;  Pred. No. 7.9e-35;
  Matches  224;  Conservative    1;  Mismatches   87;  Indels    9;  Gaps    1;

Qy         36 CAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCA 95
              | |||    || |  | ||| | |    |  ||| |       |     | ||||||||||||
Db      23848 CTAATTCCAAAATAAAATTCTTAACTCCTGCAGGGGCAGTGGCTCACGCCTGTAATCCCA 23789

Qy         96 GCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGG 155
              |||||| |||||||  ||||||| |||||||||||||||| ||| |||||||||  |||||
Db      23788 GCACTTTGGGAGGCTGAGGTGGGTGGATCACCTGAGGTCAGGAGTTCGAGACCAGCCTGG 23729

Qy        156 CCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCAC 215
              |||| |||||||||||| |||| |||||||||||||||| |||  |||| ||||||||
Db      23728 CCAATATGGTGAAACCCAGTCTCTACTAAAAATACAAAAATTAGCCGGGCGTGGTGGCAT 23669

Qy        216 ACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTG 275
              |  |||||||||||||||||||||||| | |||   | ||    |    || | | |   |
Db      23668 ATGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGAAGAATGGCTTGAACCCGGGAGG 23609

Qy        276 AGCCGAAATCACAGATCACAGAGTGAG---------CAGAGTGAGACKCCGTCTCAAAAA 326
              |.        ||  || |  ||   | |         |||||  ||||:| ||||||||||
Db      23608 CGAAAGTTGCAGTGAGCCGAGGTCGTGCCACTGCGACAGAGCAAGACTCTGTCTCAAAAA 23549

Qy        327 CAACAACAAAAAACAAAAAAA 347
              |||||||||  ||  | ||  ||
Db      23548 CAACAACAACAACAACAACAA 23528


RESULT 10
US-09-949-016-13194/c
; Sequence 13194, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13194
;   LENGTH: 152524
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13194

  Query Match             38.4%;  Score 153.4;  DB 3;  Length 152524;
  Best Local Similarity   69.8%;  Pred. No. 7.9e-35;
  Matches  224;  Conservative    1;  Mismatches   87;  Indels    9;  Gaps    1;

Qy         36 CAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCA 95
              | |||    || |  | ||| | |    |  ||| |       |     | ||||||||||||
Db      23848 CTAATTCCAAAATAAAATTCTTAACTCCTGCAGGGGCAGTGGCTCACGCCTGTAATCCCA 23789

Qy         96 GCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGG 155
              |||||| |||||||  ||||||| |||||||||||||||| ||| |||||||||  |||||
Db      23788 GCACTTTGGGAGGCTGAGGTGGGTGGATCACCTGAGGTCAGGAGTTCGAGACCAGCCTGG 23729

Qy        156 CCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCAC 215
              |||| |||||.|||||| |||| |||||||||||||||| |||| |||| ||||||||
Db      23728 CCAATATGGTGAAACCCAGTCTCTACTAAAAATACAAAAATTAGCCGGGCGTGGTGGCAT 23669

Qy        216 ACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTG 275
              |  |||||||||||||||||||||||| | |||   | ||    |    || | | |   |
Db      23668 ATGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGAAGAATGGCTTGAACCCGGGAGG 23609

Qy        276 AGCCGAAATCACAGATCACAGAGTGAG---------CAGAGTGAGACKCCGTCTCAAAAA 326
              |.        ||  || |  ||   | |         |||||  ||||:| ||||||||||
Db      23608 CGAAAGTTGCAGTGAGCCGAGGTCGTGCCACTGCGACAGAGCAAGACTCTGTCTCAAAAA 23549

Qy        327 CAACAACAAAAAACAAAAAAA 347
              |||||||||  ||  | ||  ||
Db      23548 CAACAACAACAACAACAACAA 23528


RESULT 11
US-09-949-016-16011/c
; Sequence 16011, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
```

```
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16011
;   LENGTH: 40091
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16011

  Query Match             38.2%;  Score 152.8;  DB 3;  Length 40091;
  Best Local Similarity   75.0%;  Pred. No. 6.9e-35;
  Matches  219;  Conservative    1;  Mismatches   63;  Indels    9;  Gaps    2;

Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
              ||| |||||||||||||| ||||||| ||| ||  |||||| ||||| || |||  ||
Db      30035 CCTATAATCCCAGCACTTTGGGAGGCCGAGGCAGGTAGATCACTTGAGGCCAGGAGTCCG 29976

Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
              |||||| ||||||||||||||||||| ||| ||| ||||||||| ||||||| |||||||
Db      29975 AGACCAGCCTGGCCAACATGGTGAAATCCCATCTCTACTAAAAACACAAAAATTAGCTGG 29916

Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
              ||||||||  ||| |||||| |||||||||  |||| | ||||||||||||||||||||
Db      29915 GCATGGTGGTGCACGCCTGTAATCCCAGCTACTTGGGAGGCTGAGATTGCAGTGAGCTGA 29856

Qy        264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
              |||||        | |   ||| |   || | |  | |||||| ||||:|| ||||||
Db      29855 GATCG--------TGCCACTGCACTGCAGACTG-GGCAACAGAGTAAGACTCCATCTCAA 29805

Qy        324 AAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCC 375
              ||| || || |||| |   ||| |    || | | | |   ||| ||||| |
Db      29804 AAAAAATAAAAAAACAAGCAAACAAACAAACAAACTTTAGTTCTACGGTTGC 29753


RESULT 12
US-09-949-016-69587/c
; Sequence 69587, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 69587
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-69587

  Query Match             38.0%;  Score 152;  DB 3;  Length 601;
  Best Local Similarity   74.0%;  Pred. No. 2.2e-35;
  Matches  225;  Conservative    2;  Mismatches   57;  Indels   20;  Gaps    2;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||| ||||||| ||||||||||||||||||||||||| |||
Db        487 ACGCCTGTAATCCCAGCACTTTGGGAGGCTAAGGTGGGCGGATCACCTGAGGTCAGGAGT 428

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||||||||||||||||||| |||| ||||||||||||||||| ||||
Db        427 TCGAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGC 368

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCC----GGAGATTGCAGT 256
              || |||||||||| |||||||| ||||||||||  |||| |    | ||| |    |
Db        367 CGGACATGGTGGCACGCACCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGAATCACTT 308

Qy        257 GAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCA----------------CAGAGTG 300
              || |  :|   ||||||   || |  | |  ||||||                |   |
Db        307 GAACCCRGGAGGCAGAGGTTGCAGTGAGCCAAGATCACGCCACTGCACTTCAGCCTGGGT 248

Qy        301 AGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 360
              ||||||||||:|||||||||||||||| ||  |||||  ||||||  | || |||
Db        247 GACAGAGTGAGACTCCGTCTCAAAAACAAAAAACAAAAAACAAAAAAAAACAAAACAAAA 188

Qy        361 TCCA 364
              |||
Db        187 ACCA 184


RESULT 13
US-09-949-016-12784/c
; Sequence 12784, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12784
;   LENGTH: 17050
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(17050)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12784

  Query Match             37.9%;  Score 151.4;  DB 3;  Length 17050;
  Best Local Similarity   76.4%;  Pred. No. 1.3e-34;
  Matches  201;  Conservative    1;  Mismatches   52;  Indels    9;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||| |||||||| |||||||| |||  |||||||||| |||||||| |||
Db      11937 AAGCCTGTAATTCCAGCACTTTGGGAGGCCGAGGCCGGCGGATCACTTGAGGTCAGGAGT 11878

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||||||||||||||||||| |||||| |||||||||||||||| ||||
Db      11877 TCGAGACCAGCCTGGCCAACATGGTGAAACTCCGTCTCTACTAAAAATACAAAAATTAGC 11818

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |  | ||||||| | | |||||| |||||||||| |||| ||||| |||||| ||||
Db      11817 CGCACGTGGTGGCGCGCGCCTGTAATCCCAGCTACTCGGGAGGCGGAGGTTGCAGCGAGC 11758

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||||||      |    |   |    |   |         || || ||||||:||||||
Db      11757 AGAGATCGTGCCATTGCACTCCAGTCTGGGCGA---------CACAGCGAGACTCCGTCT 11707

Qy        321 CAAAAACAACAACAAAAAACAAA 343
              |||||| || || |||||| |||
Db      11706 CAAAAAAAAAAAAAAAAAAAAAA 11684


RESULT 14
```

```
US-09-949-016-13680/c
; Sequence 13680, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13680
;   LENGTH: 17050
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(17050)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13680

  Query Match             37.9%;  Score 151.4;  DB 3;  Length 17050;
  Best Local Similarity   76.4%;  Pred. No. 1.3e-34;
  Matches  201;  Conservative    1;  Mismatches   52;  Indels    9;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||||||| |||||||||| |||||||| |||  |||||||||| |||||||| |||
Db       11937 AAGCCTGTAATTCCAGCACTTTGGGAGGCCGAGGCCGGCGGATCACTTGAGGTCAGGAGT 11878

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               ||||||||| ||||||||||||||||||||| |||||| |||||||||||||||| ||||
Db       11877 TCGAGACCAGCCTGGCCAACATGGTGAAACTCCGTCTCTACTAAAAATACAAAAATTAGC 11818

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
                 | | ||||||| | | |||||| |||||||||||| |||| ||||| |||||| ||||
Db       11817 CGCACGTGGTGGCGCGCGCCTGTAATCCCAGCTACTCGGGAGGCGGAGGTTGCAGCGAGC 11758

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               |||||||      |    |   |      |   |          || || |||||:||||||
Db       11757 AGAGATCGTGCCATTGCACTCCAGTCTGGGCGA---------CACAGCGAGACTCCGTCT 11707

Qy         321 CAAAAACAACAACAAAAAACAAA 343
               |||||| || || |||||| |||
Db       11706 CAAAAAAAAAAAAAAAAAAAAA 11684


RESULT 15
US-09-949-016-13784
; Sequence 13784, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13784
;   LENGTH: 20099
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(20099)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13784

  Query Match             37.7%;  Score 150.8;  DB 3;  Length 20099;
  Best Local Similarity   74.0%;  Pred. No. 2.1e-34;
  Matches  225;  Conservative    1;  Mismatches   58;  Indels   20;  Gaps    2;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||||||||||||||||| ||||||| |||||||||||||||||||||||||| |||
Db        8799 ACGCCTGTAATCCCAGCACTTTGGGAGGCTAAGGTGGGCGGATCACCTGAGGTCAGGAGT 8858

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               ||||||||| |||||||||||||||||||||| |||| ||||||||||||||||| ||||
Db        8859 TCGAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGC 8918

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCC----GGAGATTGCAGT 256
                || |||||||||| |||||||| |||||||||| |||| |    | ||| |    |
Db        8919 CGGACATGGTGGCACGCACCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGAATCACTT 8978

Qy         257 GAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCA----------------CAGAGTG 300
               || |   | · ||||||   || |  | |  ||||||                |  |
Db        8979 GAACCCGGGAGGCAGAGGTTGCAGTGAGCCAAGATCACGCCACTGCACTTCAGCCTGGGT 9038

Qy         301 AGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 360
                 ||||||||||:|||||||||||||||| ||  |||||  ||||||  | || |||
Db        9039 GACAGAGTGAGACTCCGTCTCAAAAACAAAAAACAAAAAACAAAAAAAAACAAAACAAAA 9098

Qy         361 TCCA 364
               ||||
Db        9099 ACCA 9102


RESULT 16
US-09-949-016-12096/c
; Sequence 12096, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12096
;   LENGTH: 19566
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12096

  Query Match             37.6%;  Score 150.4;  DB 3;  Length 19566;
  Best Local Similarity .  72.1%;  Pred. No. 2.7e-34;
  Matches  196;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||| || |||||||||  |||  |||||||||||||||| | |
Db       14763 ATGCCTGTAATCCCAGCACTTTGGAAGGCCAAGGCAGGCAGATCACCTGAGGTCAGGGGT 14704

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || ||||||| |||||||||||||||||||||||||||||| |||||||||||||||| ||||
```

```
Db      14703 TCAAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 14644

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||||||||||||  | |||||||| ||||| ||| |||| | |||  | ||
Db      14643 TGGGCATGGTGGCATGCGCCTGTAGTTCCAGCAACTTGGGAGGCTGAGGCAGGAGAATCG 14584

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              || | |||  | | |  ||  || | ||| || ||      | |   ||
Db      14583 CTTGAACCCAGGAGGTGGAGGTTGCAGTGAGCCGAGATTGTGCCATTGCACTCCATCTCA 14524

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATA 352
              ||||||  || |  |||| |||| ||  | |
Db      14523 AAAAAACTTCATCTCAAAAAAAAAGAAAGAAA 14492


RESULT 17
US-09-949-016-14114/c
; Sequence 14114, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14114
;   LENGTH: 19567
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14114

  Query Match             37.6%;  Score 150.4;  DB 3;  Length 19567;
  Best Local Similarity   72.1%;  Pred. No. 2.7e-34;
  Matches  196;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||||| || |||||||||  ||| ||||||||||||||| | |
Db      14763 ATGCCTGTAATCCCAGCACTTTGGAAGGCCAAGGCAGGCAGATCACCTGAGGTCAGGGGT 14704

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || |||||| |||||||||||||||||||||||||||||||||| ||||||||||||| ||||
Db      14703 TCAAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 14644

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||||||||  | |||||||| ||||| |||  |||| | |||  | ||
Db      14643 TGGGCATGGTGGCATGCGCCTGTAGTTCCAGCAACTTGGGAGGCTGAGGCAGGAGAATCG 14584

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               || | |||  | | |  ||  || | ||| || ||      | |   ||
Db      14583 CTTGAACCCAGGAGGTGGAGGTTGCAGTGAGCCGAGATTGTGCCATTGCACTCCATCTCA 14524

Qy         321 CAAAAACAACAACAAAAAACAAAAAAACCATA 352
               ||||||  || |  |||| |||| ||  | |
Db      14523 AAAAAACTTCATCTCAAAAAAAAAGAAAGAAA 14492


RESULT 18
US-09-949-016-14157
; Sequence 14157, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14157
;   LENGTH: 455726
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(455726)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14157

  Query Match             37.4%;  Score 149.6;  DB 3;  Length 455726;
  Best Local Similarity   73.5%;  Pred. No. 1.6e-33;
  Matches  219;  Conservative    1;  Mismatches   69;  Indels    9;  Gaps    2;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||| ||||||| |||||||| |||||||||||||||| |||| ||| |||
Db     143833 ATGCCTGTAATCCTAGCACTTTGGGAGGCCGAGGTGGGCGGATCACTTGAGATCAGGAGT 143892

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || ||||||  |||||||||||||||||||||| ||| ||||||||||| |||   ||||
Db     143893 TCAAGACCAGTCTGGCCAACATGGTGAAACCCCATCTGTACTAAAAATATAAAGTTTAGC 143952

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||  || |||  ||   ||||| |||||||||||| ||||  |||| ||||||||||
Db     143953 TGGGTGTGATGGTGCATGTCTGTAATCCCAGCTACTCGGGAGGTGGAGGTTGCAGTGAGT 144012

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               ||||||||           | |   |||    | | | |||| ||||||||||:  ||||
Db     144013 TGAGATCG--------TGCCACTGCACTCCACCCTGGGTGA-CAGAGTGAGACTTTGTCT 144063

Qy         321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAG 378
               ||||||| || |||||| ||  |||   ||   || | |   ||   | | ||| |
Db     144064 TTAAAACAAAAAAAAAAAAAAAAGCAAATTCTACCTCACTTTTAGTCCCAGTTGCCCTG 144121


RESULT 19
US-09-949-016-11940
; Sequence 11940, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11940
;   LENGTH: 481115
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(481115)
;   OTHER INFORMATION: n = A,T,C or G
```

```
US-09-949-016-11940

  Query Match             37.4%;  Score 149.6;  DB 3;  Length 481115;
  Best Local Similarity   73.5%;  Pred. No. 1.7e-33;
  Matches  219;  Conservative    1;  Mismatches   69;  Indels    9;  Gaps    2;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||| ||||||| |||||||| |||||||||||||| |||| ||| |||
Db     121222 ATGCCTGTAATCCTAGCACTTTGGGAGGCCGAGGTGGGCGGATCACTTGAGATCAGGAGT 121281

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || ||||||  |||||||||||||||||||||| ||| ||||||||||| |||   ||||
Db     121282 TCAAGACCAGTCTGGCCAACATGGTGAAACCCCATCTGTACTAAAAATATAAAGTTTAGC 121341

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||  || |||  ||   ||||| |||||||||||| ||||  |||| ||||||||||
Db     121342 TGGGTGTGATGGTGCATGTCTGTAATCCCAGCTACTCGGGAGGTGGAGGTTGCAGTGAGT 121401

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              ||||||||          | |   |||    | | | |||| ||||||||||:  ||||
Db     121402 TGAGATCG--------TGCCACTGCACTCCACCCTGGGTGA-CAGAGTGAGACTTTGTCT 121452

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAG 378
                |||||| || |||||| ||  |||   ||   || | |   ||   | | ||| |
Db     121453 TTAAAACAAAAAAAAAAAAAAAGCAAATTCTACCTCACTTTTAGTCCCAGTTGCCCTG 121510


RESULT 20
US-09-949-016-16723/c
; Sequence 16723, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16723
;   LENGTH: 157822
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16723

  Query Match             37.4%;  Score 149.4;  DB 3;  Length 157822;
  Best Local Similarity   74.0%;  Pred. No. 1.2e-33;
  Matches  205;  Conservative    1;  Mismatches   62;  Indels    9;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||||| |||||||  |||    | | |||||||  | |||| |||
Db      15027 ACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAAGTGCATCACCTAAAGTCACGAGT 14968

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || |||||| |||||||||||||   |||| | |||| | ||||||||||||||| ||||
Db      14967 TCAAGACCAGCCTGGCCAACATGGCAAAACTCTGTCTCTGCTAAAAATACAAAAATTAGC 14908

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
                ||||| ||||||  |||||||| |||||||||||| |||| |||||  ||||||||||
Db      14907 CAGGCATAGTGGCAGGCACCTGTAATCCCAGCTACTCTGGAGGCGGAGGTTGCAGTGAGC 14848

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||||| |||         |   | |   | | |          |||||  ||||:| ||||
Db      14847 TGAGATTGCACCACTGCTCTTTAGCCTGGGTGA---------CAGAGCAAGACTCTGTCT 14797

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
                ||||||| || || |||||| ||  |||  | || |||
Db      14796 CAAAAAAAAAAAAAAAAAAAAAGCAAAAAACAAAACA 14760


RESULT 21
US-09-949-016-17127
; Sequence 17127, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17127
;   LENGTH: 228896
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(228896)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17127

  Query Match             37.3%;  Score 149;  DB 3;  Length 228896;
  Best Local Similarity   68.4%;  Pred. No. 1.9e-33;
  Matches  236;  Conservative    1;  Mismatches   76;  Indels   32;  Gaps    1;

Qy         46 TAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGG 105
              ||  | ||    ||||  | || |     |    | |||||||||||||||||| |||
Db      55811 TAGAAAATGACAAGGCTGGCCAGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGG 55870

Qy        106 AGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGT 165
              ||||| |||||||||  |   ||||| |||| ||| |||||||||| |||||| ||||||||
Db      55871 AGGCCGAGGTGGGCGAACTGCCTGACGTCAGGAGTTCGAGACCAGCCTGGCTAACATGGT 55930

Qy        166 GAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAG 225
              |||||||||||| |||||||||||||||||||| ||||||||| ||||||||||||||||||
Db      55931 GAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCGTGGTGGCACACACCTGTAG 55990

Qy        226 TCCCAGCT----------------------------------ACTCAGGAGCCGGAGATTGC 253
              ||||||||                                  ||||||||| | ||| ||||
Db      55991 TCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATTGCTTGAACTCAGGAGGCAGAGGTTGC 56050

Qy        254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGAC 313
              |||||||||||||| ||         |   | |   ||      | | | ||||||  ||||
Db      56051 AGTGAGCTGAGATCACACCACTGCACTCCAGCCTGGAGGGGGGTGGGGGCAGAGCAAGAC 56110

Qy        314 KCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACAT 358
              :| |||||||||||| |||||||| | || ||  | |  | ||
Db      56111 TCTGTCTCAAAAACAAAAACAAAAAAAACAACAAAAAAAGAAAAT 56155


RESULT 22
US-09-949-016-12517/c
; Sequence 12517, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12517
;   LENGTH: 77994
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12517

  Query Match             37.2%;  Score 148.8;  DB 3;  Length 77994;
  Best Local Similarity   72.7%;  Pred. No. 1.4e-33;
  Matches  192;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
              |||||||||||||||||| |||||||||||| ||| |  |||| |||||||| ||| |||
Db      73058 CCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGTGATTCACTTGAGGTCAGGAGTTCG 72999

Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
              |||||| |||| | ||||||||||||||||||||||||||||||||||||||| |||||||
Db      72998 AGACCAGCCTGACTAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAATTAGCTGG 72939

Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
              || ||||||| |||| ||||| |||||||||||  |||| | |||     ||
Db      72938 GCGTGGTGGCGCACATCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAAGAGAATTGCAT 72879

Qy        264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
              || | |||| | | |   ||| |  | |     |       | |||   |  | ||
Db      72878 GAACCCAGAAGGCGGAGGTTGCACTGCACTCTAGCTTGAGCAACAGAGCAACTCTGTCTC 72819

Qy        324 AAACAACAACAAAAAACAAAAAAA 347
              ||| || || |||||| |||||||
Db      72818 AAAAAAAAAAAAAAAAAAAAAAA 72795


RESULT 23
US-09-949-016-16021/c
; Sequence 16021, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16021
;   LENGTH: 77994
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16021

  Query Match             37.2%;  Score 148.8;  DB 3;  Length 77994;
  Best Local Similarity   72.7%;  Pred. No. 1.4e-33;
  Matches  192;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
              |||||||||||||||||| |||||||||||| ||| |  |||| |||||||| ||| |||
Db      73058 CCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGTGATTCACTTGAGGTCAGGAGTTCG 72999

Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
              |||||| |||| | ||||||||||||||||||||||||||||||||||||||| |||||||
Db      72998 AGACCAGCCTGACTAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAATTAGCTGG 72939

Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
              || ||||||| |||| ||||| |||||||||||  |||| | |||     ||
Db      72938 GCGTGGTGGCGCACATCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAAGAGAATTGCAT 72879

Qy        264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
              || | |||| | | |   ||| |  | |     |       | |||   |  | ||
Db      72878 GAACCCAGAAGGCGGAGGTTGCACTGCACTCTAGCTTGAGCAACAGAGCAACTCTGTCTC 72819

Qy        324 AAACAACAACAAAAAACAAAAAAA 347
              ||| || || |||||| |||||||
Db      72818 AAAAAAAAAAAAAAAAAAAAAAA 72795


RESULT 24
US-09-949-016-13460/c
; Sequence 13460, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13460
;   LENGTH: 8100
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13460

  Query Match             37.2%;  Score 148.6;  DB 3;  Length 8100;
  Best Local Similarity   74.9%;  Pred. No. 6.4e-34;
  Matches  200;  Conservative    1;  Mismatches   60;  Indels    6;  Gaps    1;

Qy         83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
              |||||||||||||||||| |||||||||||| || |||||||||||||||    | ||
Db       6648 GCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTAGGTGGATCACCTGAGGTCAGAGGTTC 6589

Qy        143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
              ||||||  |||||||||||||||||||||| || |||||||||||||||| ||||||
Db       6588 AAGACCAGTCTGGCCAACATGGTGAAACCCCGCCTCTACTAAAAATACAAAAATTAGCTG 6529

Qy        203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
              |||||||||||  | |||||| ||||||||||  |||| | |||   | ||
Db       6528 GGCATGGTGGCAGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCT 6469

Qy        263 AGATCGCAGAGTGAGCCGAAATCACAGATC------ACAGAGTGAGCAGAGTGAGACKCC 316
              || | |||   |     |    ||  || |      ||   |  | ||||||||||:|
Db       6468 TGAACCCAGGAGGCAGAGGTTGCAGTGAGCACTCCAACCTGGGCAACAGAGTGAGACTCT 6409

Qy        317 GTCTCAAAAACAACAACAAAAAACAAA 343
              |||||||||| || || |||||| |||
Db       6408 GTCTCAAAAAAAAAAAAAAAAAAAAA 6382


RESULT 25
US-09-949-016-13461/c
; Sequence 13461, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13461
; LENGTH: 8100
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13461

Query Match 37.2%; Score 148.6; DB 3; Length 8100;
Best Local Similarity 74.9%; Pred. No. 6.4e-34;
Matches 200; Conservative 1; Mismatches 60; Indels 6; Gaps 1;

Qy 83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
||||||||||||||||||||| |||||||||||||| || |||||||||||||||| | ||
Db 6648 GCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTAGGTGGATCACCTGAGGTCAGAGGTTC 6589

Qy 143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
|||||| |||||||||||||||||||||| || |||||||||||||||| ||||||
Db 6588 AAGACCAGTCTGGCCAACATGGTGAAACCCCGCCTCTACTAAAAATACAAAAATTAGCTG 6529

Qy 203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
|||||||||||| | |||||| |||||||||||| |||| | ||| | ||
Db 6528 GGCATGGTGGCAGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCT 6469

Qy 263 AGATCGCAGAGTGAGCCGAAATCACAGATC------ACAGAGTGAGCAGAGTGAGACKCC 316
|| | ||| | | || || | || | | |||||||||||:|
Db 6468 TGAACCCAGGAGGCAGAGGTTGCAGTGAGCACTCCAACCTGGGCAACAGAGTGAGACTCT 6409

Qy 317 GTCTCAAAAACAACAACAAAAAACAAA 343
|||||||||| || || |||||| |||
Db 6408 GTCTCAAAAAAAAAAAAAAAAAAAAAA 6382


RESULT 26
US-09-949-016-152016
; Sequence 152016, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 152016
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-152016

Query Match 37.0%; Score 148; DB 3; Length 601;
Best Local Similarity 75.0%; Pred. No. 3.4e-34;
Matches 213; Conservative 1; Mismatches 61; Indels 9; Gaps 2;

Qy 84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
||| |||||||||||||| |||||||| ||| || |||||| ||||| || ||| ||
Db 326 CCTATAATCCCAGCACTTTGGGAGGCCGAGGCAGGTAGATCACTTGAGGCCAGGAGTCCG 385

Qy 144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
|||||| |||||||||||||||||||| ||| ||| ||||||||| ||||||| |||||||
Db 386 AGACCAGCCTGGCCAACATGGTGAAATCCCATCTCTACTAAAAACACAAAAATTAGCTGG 445

Qy 204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
||||||||| ||| |||||| |||||||||| |||| | ||||||||||||||||||
Db 446 GCATGGTGGTGCACGCCTGTAATCCCAGCTACTTGGGAGGCTGAGATTGCAGTGAGCTGA 505

Qy 264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
||||| | | ||| | || | | | ||||||| ||||:|| ||||||
Db 506 GATCG--------TGCCACTGCACTGCAGACTG-GGCAACAGAGTAAGACTCCATCTCAA 556

Qy 324 AAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCT 367
||| || || |||| | ||| | || | | | |||
Db 557 AAAAAATAAAAAAACAAGCAAACAAACAAACAAACTTTAGTTCT 600


RESULT 27
US-09-949-002-2953/c
; Sequence 2953, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2953
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-2953

Query Match 37.0%; Score 147.8; DB 3; Length 601;
Best Local Similarity 78.9%; Pred. No. 3.9e-34;
Matches 176; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
|||||||||||||||||||| |||||||| ||||||| ||||| ||||||||||| |||
Db 232 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCGCCTGAGGTCAGGAGT 173

Qy 141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
| |||||| |||||||||||||||||||||||| |||||||||||||||| ||||
Db 172 TTGAGACCGGCCTGGCCAACATGGTGAAACCCCGTCACTACTAAAAATACAAAAATTAGC 113

Qy 201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
|||||||||||||||||| |||||||||||||||||||| | ||| ||
Db 112 TGGGCATGGTGGCACACACTTGTAGTCCCAGCTACTCAGGAGGCTGAGGCATGAGAATTG 53

Qy 261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
|| | |||| | | | || || | ||| | ||
Db 52 CTTGAACCCAGAAGGTGGAGGTTGCAGTGAGCCAAGATGGTGC 10


RESULT 28
US-09-949-002-2954/c
; Sequence 2954, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
```

```
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE:  2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2954
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-2954

  Query Match             37.0%;  Score 147.8;  DB 3;  Length 601;
  Best Local Similarity   78.9%;  Pred. No. 3.9e-34;
  Matches  176;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||| ||||||| ||||| |||||||||| |||
Db        253 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCGCCTGAGGTCAGGAGT 194

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||  |||||||||||||||||||||||||  |||||||||||||||| ||||
Db        193 TTGAGACCGGCCTGGCCAACATGGTGAAACCCCGTCACTACTAAAAATACAAAAATTAGC 134

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||||||||||||| |||||||||||||||||||| | |||      ||
Db        133 TGGGCATGGTGGCACACACTTGTAGTCCCAGCTACTCAGGAGGCTGAGGCATGAGAATTG 74

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
               || | ||||  | |  |    ||  || |  |||  | ||
Db         73 CTTGAACCCAGAAGGTGGAGGTTGCAGTGAGCCAAGATGGTGC 31


RESULT 29
US-09-949-002-4465/c
; Sequence 4465, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4465
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-4465

  Query Match             37.0%;  Score 147.8;  DB 3;  Length 601;
  Best Local Similarity   78.9%;  Pred. No. 3.9e-34;
  Matches  176;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||| ||||||| ||||| |||||||||| |||
Db        232 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCGCCTGAGGTCAGGAGT 173

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||  |||||||||||||||||||||||||  |||||||||||||||| ||||
Db        172 TTGAGACCGGCCTGGCCAACATGGTGAAACCCCGTCACTACTAAAAATACAAAAATTAGC 113

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||||||||||||| |||||||||||||||||||| | |||      ||
Db        112 TGGGCATGGTGGCACACACTTGTAGTCCCAGCTACTCAGGAGGCTGAGGCATGAGAATTG 53

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
               || | ||||  | |  |    ||  || |  |||  | ||
Db         52 CTTGAACCCAGAAGGTGGAGGTTGCAGTGAGCCAAGATGGTGC 10


RESULT 30
US-09-949-002-4466/c
; Sequence 4466, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4466
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-4466

  Query Match             37.0%;  Score 147.8;  DB 3;  Length 601;
  Best Local Similarity   78.9%;  Pred. No. 3.9e-34;
  Matches  176;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||| ||||||| ||||| |||||||||| |||
Db        253 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCGCCTGAGGTCAGGAGT 194

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||  |||||||||||||||||||||||||  |||||||||||||||| ||||
Db        193 TTGAGACCGGCCTGGCCAACATGGTGAAACCCCGTCACTACTAAAAATACAAAAATTAGC 134

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||||||||||||| |||||||||||||||||||| | |||      ||
Db        133 TGGGCATGGTGGCACACACTTGTAGTCCCAGCTACTCAGGAGGCTGAGGCATGAGAATTG 74

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
               || | ||||  | |  |    ||  || |  |||  | ||
Db         73 CTTGAACCCAGAAGGTGGAGGTTGCAGTGAGCCAAGATGGTGC 31


RESULT 31
US-09-949-002-650
; Sequence 650, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 650
;   LENGTH: 60595
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-650

  Query Match             37.0%;  Score 147.8;  DB 3;  Length 60595;
  Best Local Similarity   78.9%;  Pred. No. 2.5e-33;
  Matches  176;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
```

```
                  ||||||||||||||||||||| |||||||| ||||||| ||||| |||||||||| |||
Db      55522 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCGCCTGAGGTCAGGAGT 55581

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | |||||||  |||||||||||||||||||||||||||||||  |||||||||||||||| ||||
Db      55582 TTGAGACCGGCCTGGCCAACATGGTGAAACCCCGTCACTACTAAAAATACAAAAATTAGC 55641

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||||||||||||||||||||||||||||||||||||| | |||      ||
Db      55642 TGGGCATGGTGGCACACACTTGTAGTCCCAGCTACTCAGGAGGCTGAGGCATGAGAATTG 55701

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
                || | ||||  | |  |     ||  || |  |||  | ||
Db      55702 CTTGAACCCAGAAGGTGGAGGTTGCAGTGAGCCAAGATGGTGC 55744


RESULT 32
US-09-949-002-706
; Sequence 706, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 706
;   LENGTH: 60595
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-706

  Query Match           37.0%;  Score 147.8;  DB 3;  Length 60595;
  Best Local Similarity   78.9%;  Pred. No. 2.5e-33;
  Matches  176;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||||| |||||||| ||||||| ||||| |||||||||| |||
Db      55522 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCGCCTGAGGTCAGGAGT 55581

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | |||||||  |||||||||||||||||||||||||||||||  |||||||||||||||| ||||
Db      55582 TTGAGACCGGCCTGGCCAACATGGTGAAACCCCGTCACTACTAAAAATACAAAAATTAGC 55641

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||||||||||||||||||||||||||||||||||||| | |||      ||
Db      55642 TGGGCATGGTGGCACACACTTGTAGTCCCAGCTACTCAGGAGGCTGAGGCATGAGAATTG 55701

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
                || | ||||  | |  |     ||  || |  |||  | ||
Db      55702 CTTGAACCCAGAAGGTGGAGGTTGCAGTGAGCCAAGATGGTGC 55744


RESULT 33
US-09-949-016-13849
; Sequence 13849, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13849
;   LENGTH: 86945
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(86945)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13849

  Query Match           37.0%;  Score 147.8;  DB 3;  Length 86945;
  Best Local Similarity   73.5%;  Pred. No. 2.9e-33;
  Matches  208;  Conservative    1;  Mismatches   58;  Indels   16;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||||| |||||||  |||||||| |||| || |||||||| |||
Db      25003 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGAGGATTACTTGAGGTCAGGAGT 25062

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | |||||||  |||||||||||||||| ||||||  |||| ||||||||||||||||  |||
Db      25063 TTGAGACCAGCCTGGCCAACATGGTAAAACCCTGTCTCTACTAAAAATACAAAAATGAGC 25122

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||  ||||||     ||||| || ||| |||||||||||| | ||| ||||||||| |
Db      25123 TGGGCGTGGTGGTGTGCACCTATAATCCTAGCTACTCAGGAGGCTGAGGTTGCAGTGAAC 25182

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGAT----------------CACAGAGTGAGCA 304
                |   ||||||   || |  | |  ||||                |   | | ||
Db      25183 CTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCCTGGGCAACA 25242

Qy        305 GAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
              | |||||||:|| |||||||| || || | |||| |||| ||
Db      25243 GTGTGAGACTCCATCTCAAAAAGAATAAAATAAAATAAAATAA 25285


RESULT 34
US-09-949-016-13850
; Sequence 13850, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13850
;   LENGTH: 86945
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(86945)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13850

  Query Match           37.0%;  Score 147.8;  DB 3;  Length 86945;
  Best Local Similarity   73.5%;  Pred. No. 2.9e-33;
  Matches  208;  Conservative    1;  Mismatches   58;  Indels   16;  Gaps    1;
```

```
Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               ||||||||||||||||||||| |||||||  ||||||| |||| || ||||||||  |||
Db       25003 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGAGGATTACTTGAGGTCAGGAGT 25062

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               | ||||||| |||||||||||||||| |||||| |||| ||||||||||||||||||  |||
Db       25063 TTGAGACCAGCCTGGCCAACATGGTAAAACCCTGTCTCTACTAAAAATACAAAAATGAGC 25122

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||| ||||||    ||||| || ||| |||||||||||||| | ||| |||||||||  |
Db       25123 TGGGCGTGGTGGTGTGCACCTATAATCCTAGCTACTCAGGAGGCTGAGGTTGCAGTGAAC 25182

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGAT----------------CACAGAGTGAGCA 304
               |   ||||||  || |  | |  ||||                  |   |  | ||
Db       25183 CTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCCTGGGCAACA 25242

Qy         305 GAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
               | |||||||:|| ||||||||| || || | |||| |||| ||
Db       25243 GTGTGAGACTCCATCTCAAAAAGAATAAAATAAAATAAAATAA 25285


RESULT 35
US-09-949-016-13851
; Sequence 13851, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13851
;   LENGTH: 86945
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(86945)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13851

  Query Match             37.0%;  Score 147.8;  DB 3;  Length 86945;
  Best Local Similarity   73.5%;  Pred. No. 2.9e-33;
  Matches  208;  Conservative    1;  Mismatches   58;  Indels   16;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||| |||||||  ||||||| |||| || ||||||||  |||
Db       25003 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGAGGATTACTTGAGGTCAGGAGT 25062

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               | ||||||| |||||||||||||||| |||||| |||| ||||||||||||||||||  |||
Db       25063 TTGAGACCAGCCTGGCCAACATGGTAAAACCCTGTCTCTACTAAAAATACAAAAATGAGC 25122

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||| ||||||    ||||| || ||| |||||||||||||| | ||| |||||||||  |
Db       25123 TGGGCGTGGTGGTGTGCACCTATAATCCTAGCTACTCAGGAGGCTGAGGTTGCAGTGAAC 25182

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGAT----------------CACAGAGTGAGCA 304
               |   ||||||  || |  | |  ||||                  |   |  | ||
Db       25183 CTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCCTGGGCAACA 25242

Qy         305 GAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
               | |||||||:|| ||||||||| || || | |||| |||| ||
Db       25243 GTGTGAGACTCCATCTCAAAAAGAATAAAATAAAATAAAATAA 25285


RESULT 36
US-09-949-016-13852
; Sequence 13852, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13852
;   LENGTH: 86945
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(86945)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13852

  Query Match             37.0%;  Score 147.8;  DB 3;  Length 86945;
  Best Local Similarity   73.5%;  Pred. No. 2.9e-33;
  Matches  208;  Conservative    1;  Mismatches   58;  Indels   16;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||| |||||||  ||||||| |||| || ||||||||  |||
Db       25003 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGAGGATTACTTGAGGTCAGGAGT 25062

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               | ||||||| |||||||||||||||| |||||| |||| ||||||||||||||||||  |||
Db       25063 TTGAGACCAGCCTGGCCAACATGGTAAAACCCTGTCTCTACTAAAAATACAAAAATGAGC 25122

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||| ||||||    ||||| || ||| |||||||||||||| | ||| |||||||||  |
Db       25123 TGGGCGTGGTGGTGTGCACCTATAATCCTAGCTACTCAGGAGGCTGAGGTTGCAGTGAAC 25182

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGAT----------------CACAGAGTGAGCA 304
               |   ||||||  || |  | |  ||||                  |   |  | ||
Db       25183 CTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCCTGGGCAACA 25242

Qy         305 GAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
               | |||||||:|| ||||||||| || || | |||| |||| ||
Db       25243 GTGTGAGACTCCATCTCAAAAAGAATAAAATAAAATAAAATAA 25285


RESULT 37
US-09-949-016-13853
; Sequence 13853, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
```

```
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13853
;   LENGTH: 86945
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(86945)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13853

  Query Match            37.0%;  Score 147.8;  DB 3;  Length 86945;
  Best Local Similarity   73.5%;  Pred. No. 2.9e-33;
  Matches  208;  Conservative    1;  Mismatches   58;  Indels   16;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||  |||||||  |||| || ||||||||  |||
Db      25003 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGAGGATTACTTGAGGTCAGGAGT 25062

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              | ||||||| |||||||||||||||| |||||| |||| ||||||||||||||||||  |||
Db      25063 TTGAGACCAGCCTGGCCAACATGGTAAAACCCTGTCTCTACTAAAAATACAAAAATGAGC 25122

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||| ||||||    ||||| || ||| |||||||||||||| | ||| |||||||| |
Db      25123 TGGGCGTGGTGGTGTGCACCTATAATCCTAGCTACTCAGGAGGCTGAGGTTGCAGTGAAC 25182

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGAT----------------CACAGAGTGAGCA 304
                |  ||||||   || |  | |  ||||                |   |  | ||
Db      25183 CTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCCTGGGCAACA 25242

Qy        305 GAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
              | |||||||:|| |||||||| || || | |||| |||| ||
Db      25243 GTGTGAGACTCCATCTCAAAAAGAATAAAATAAAATAAAATAA 25285


RESULT 38
US-09-949-016-13854
; Sequence 13854, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13854
;   LENGTH: 86945
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(86945)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13854

  Query Match            37.0%;  Score 147.8;  DB 3;  Length 86945;
  Best Local Similarity   73.5%;  Pred. No. 2.9e-33;
  Matches  208;  Conservative    1;  Mismatches   58;  Indels   16;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||  |||||||  |||| || ||||||||  |||
Db      25003 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGAGGATTACTTGAGGTCAGGAGT 25062

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              | ||||||| |||||||||||||||| |||||| |||| ||||||||||||||||||  |||
Db      25063 TTGAGACCAGCCTGGCCAACATGGTAAAACCCTGTCTCTACTAAAAATACAAAAATGAGC 25122

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||| ||||||    ||||| || ||| |||||||||||||| | ||| |||||||| |
Db      25123 TGGGCGTGGTGGTGTGCACCTATAATCCTAGCTACTCAGGAGGCTGAGGTTGCAGTGAAC 25182

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGAT----------------CACAGAGTGAGCA 304
                |  ||||||   || |  | |  ||||                |   |  | ||
Db      25183 CTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCCTGGGCAACA 25242

Qy        305 GAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
              | |||||||:|| |||||||| || || | |||| |||| ||
Db      25243 GTGTGAGACTCCATCTCAAAAAGAATAAAATAAAATAAAATAA 25285


Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              | ||||||| |||||||||||||||| |||||| |||| ||||||||||||||||||  |||
Db      25063 TTGAGACCAGCCTGGCCAACATGGTAAAACCCTGTCTCTACTAAAAATACAAAAATGAGC 25122

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||| ||||||    ||||| || ||| |||||||||||||| | ||| |||||||| |
Db      25123 TGGGCGTGGTGGTGTGCACCTATAATCCTAGCTACTCAGGAGGCTGAGGTTGCAGTGAAC 25182

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGAT----------------CACAGAGTGAGCA 304
                |  ||||||   || |  | |  ||||                |   |  | ||
Db      25183 CTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCCTGGGCAACA 25242

Qy        305 GAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
              | |||||||:|| |||||||| || || | |||| |||| ||
Db      25243 GTGTGAGACTCCATCTCAAAAAGAATAAAATAAAATAAAATAA 25285


RESULT 39
US-09-949-016-13855
; Sequence 13855, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13855
;   LENGTH: 86945
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(86945)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13855

  Query Match            37.0%;  Score 147.8;  DB 3;  Length 86945;
  Best Local Similarity   73.5%;  Pred. No. 2.9e-33;
  Matches  208;  Conservative    1;  Mismatches   58;  Indels   16;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||  |||||||  |||| || ||||||||  |||
Db      25003 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGAGGATTACTTGAGGTCAGGAGT 25062

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              | ||||||| |||||||||||||||| |||||| |||| ||||||||||||||||||  |||
Db      25063 TTGAGACCAGCCTGGCCAACATGGTAAAACCCTGTCTCTACTAAAAATACAAAAATGAGC 25122

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||| ||||||    ||||| || ||| |||||||||||||| | ||| |||||||| |
Db      25123 TGGGCGTGGTGGTGTGCACCTATAATCCTAGCTACTCAGGAGGCTGAGGTTGCAGTGAAC 25182

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGAT----------------CACAGAGTGAGCA 304
                |  ||||||   || |  | |  ||||                |   |  | ||
Db      25183 CTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCCTGGGCAACA 25242

Qy        305 GAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
              | |||||||:|| |||||||| || || | |||| |||| ||
Db      25243 GTGTGAGACTCCATCTCAAAAAGAATAAAATAAAATAAAATAA 25285


RESULT 40
```

```
US-09-949-016-13856
; Sequence 13856, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13856
;   LENGTH: 86945
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(86945)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13856

  Query Match             37.0%;  Score 147.8;  DB 3;  Length 86945;
  Best Local Similarity   73.5%;  Pred. No. 2.9e-33;
  Matches  208;  Conservative    1;  Mismatches   58;  Indels   16;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||||| |||||||  ||||||| |||| || ||||||||  |||
Db      25003 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGAGGATTACTTGAGGTCAGGAGT 25062

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||| |||||||||||||||| |||||| |||| ||||||||||||||||  |||
Db      25063 TTGAGACCAGCCTGGCCAACATGGTAAAACCCTGTCTCTACTAAAAATACAAAAATGAGC 25122

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||| ||||||    ||||| || ||| |||||||||||||| | ||| ||||||||| |
Db      25123 TGGGCGTGGTGGTGTGCACCTATAATCCTAGCTACTCAGGAGGCTGAGGTTGCAGTGAAC 25182

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGAT----------------CACAGAGTGAGCA 304
                |   ||||||   || |  | |  ||||                |   |  | ||
Db      25183 CTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCCTGGGCAACA 25242

Qy        305 GAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
              | |||||||:|| ||||||||| || || | |||| |||| ||
Db      25243 GTGTGAGACTCCATCTCAAAAAGAATAAAATAAAATAAAATAA 25285


RESULT 41
US-09-949-016-13857
; Sequence 13857, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13857
;   LENGTH: 86945
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(86945)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13857

  Query Match             37.0%;  Score 147.8;  DB 3;  Length 86945;
  Best Local Similarity   73.5%;  Pred. No. 2.9e-33;
  Matches  208;  Conservative    1;  Mismatches   58;  Indels   16;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||||| |||||||  ||||||| |||| || ||||||||  |||
Db      25003 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGAGGATTACTTGAGGTCAGGAGT 25062

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||| |||||||||||||||| |||||| |||| ||||||||||||||||  |||
Db      25063 TTGAGACCAGCCTGGCCAACATGGTAAAACCCTGTCTCTACTAAAAATACAAAAATGAGC 25122

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||| ||||||    ||||| || ||| |||||||||||||| | ||| ||||||||| |
Db      25123 TGGGCGTGGTGGTGTGCACCTATAATCCTAGCTACTCAGGAGGCTGAGGTTGCAGTGAAC 25182

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGAT----------------CACAGAGTGAGCA 304
                |   ||||||   || |  | |  ||||                |   |  | ||
Db      25183 CTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCCTGGGCAACA 25242

Qy        305 GAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
              | |||||||:|| ||||||||| || || | |||| |||| ||
Db      25243 GTGTGAGACTCCATCTCAAAAAGAATAAAATAAAATAAAATAA 25285


RESULT 42
US-09-949-016-13858
; Sequence 13858, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13858
;   LENGTH: 86945
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(86945)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13858

  Query Match             37.0%;  Score 147.8;  DB 3;  Length 86945;
  Best Local Similarity   73.5%;  Pred. No. 2.9e-33;
  Matches  208;  Conservative    1;  Mismatches   58;  Indels   16;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||||| |||||||  ||||||| |||| || ||||||||  |||
Db      25003 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGAGGATTACTTGAGGTCAGGAGT 25062

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||| |||||||||||||||| |||||| |||| ||||||||||||||||  |||
```

```
Db      25063 TTGAGACCAGCCTGGCCAACATGGTAAAACCCTGTCTCTACTAAAAATACAAAAATGAGC 25122

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260

Db      25123 TGGGCGTGGTGGTGTGCACCTATAATCCTAGCTACTCAGGAGGCTGAGGTTGCAGTGAAC 25182

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGAT----------------CACAGAGTGAGCA 304

Db      25183 CTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCCTGGGCAACA 25242

Qy        305 GAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347

Db      25243 GTGTGAGACTCCATCTCAAAAAGAATAAAATAAAATAAAATAA 25285


RESULT 43
US-09-949-016-14022
; Sequence 14022, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14022
;   LENGTH: 9798
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14022

  Query Match             36.9%;  Score 147.6;  DB 3;  Length 9798;
  Best Local Similarity   75.4%;  Pred. No. 1.4e-33;
  Matches  199;  Conservative    1;  Mismatches   55;  Indels    9;  Gaps    1;

Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143

Db        964 CCTGTAATCCCGGCACTTTGGGAAGCCAAGGCAGGCAGGTTGCTTGAGGCCAGGAGTTCA 1023

Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203

Db       1024 AGACCAGCCTGGCCCACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCGG 1083

Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263

Db       1084 ACATGGTGGCACGTGCCTGTAATCCCAGGTACTCAAGAGGTGGAGGTTGCAGTGAGCGGA 1143

Qy        264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323

Db       1144 GATCACAGCCCTGCACTCCAGCCTGGGTGA---------CAGAGCGAGACTCCATCTCAA 1194

Qy        324 AAACAACAACAAAAAACAAAAAAA 347

Db       1195 AACGAAAAACAAAAAACAGCACAA 1218


RESULT 44
US-09-949-016-12819
; Sequence 12819, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12819
;   LENGTH: 9801
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12819

  Query Match             36.9%;  Score 147.6;  DB 3;  Length 9801;
  Best Local Similarity   75.4%;  Pred. No. 1.4e-33;
  Matches  199;  Conservative    1;  Mismatches   55;  Indels    9;  Gaps    1;

Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143

Db        964 CCTGTAATCCCGGCACTTTGGGAAGCCAAGGCAGGCAGGTTGCTTGAGGCCAGGAGTTCA 1023

Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203

Db       1024 AGACCAGCCTGGCCCACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCGG 1083

Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263

Db       1084 ACATGGTGGCACGTGCCTGTAATCCCAGGTACTCAAGAGGTGGAGGTTGCAGTGAGCGGA 1143

Qy        264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323

Db       1144 GATCACAGCCCTGCACTCCAGCCTGGGTGA---------CAGAGCGAGACTCCATCTCAA 1194

Qy        324 AAACAACAACAAAAAACAAAAAAA 347

Db       1195 AACGAAAAACAAAAAACAGCACAA 1218


RESULT 45
US-09-936-271C-56
; Sequence 56, Application US/09936271C
; Patent No. 7022497
; GENERAL INFORMATION:
;  APPLICANT: Yousef, George M.
;  APPLICANT:  Diamandis, Eleftherios
;  TITLE OF INVENTION: No. 7022497el Human Kallikrein-Like Genes
;  FILE REFERENCE: MTS3USA
;  CURRENT APPLICATION NUMBER: US/09/936,271C
;  CURRENT FILING DATE:  2001-09-10
;  PRIOR APPLICATION NUMBER: PCT/CA00/00258
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: US 60/124,260
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/127,386
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: US 60/144,919
;  PRIOR FILING DATE: 1999-07-21
;  NUMBER OF SEQ ID NOS: 97
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
;   LENGTH: 11820
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-936-271C-56

  Query Match             36.9%;  Score 147.6;  DB 5;  Length 11820;
  Best Local Similarity   75.4%;  Pred. No. 1.5e-33;
  Matches  199;  Conservative    1;  Mismatches   55;  Indels    9;  Gaps    1;
```

```
Qy          84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
               |||||||||||| |||||| |||| |||||||  ||| | |  | ||||| || ||| ||
Db        2706 CCTGTAATCCCGGCACTTTGGGAAGCCAAGGCAGGCAGGTTGCTTGAGGCCAGGAGTTCA 2765

Qy         144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
               |||||| ||||||| |||||||||||||||||| |||||||||||||||| |||| ||
Db        2766 AGACCAGCCTGGCCCACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCGG 2825

Qy         204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
               |||||||||||   |||||| |||||| |||||| |||  |||| |||||||||||| ||
Db        2826 ACATGGTGGCACGTGCCTGTAATCCCAGGTACTCAAGAGGTGGAGGTTGCAGTGAGCGGA 2885

Qy         264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
               |||| |||       |  | |  | | |          ||||| |||||:|| ||||||
Db        2886 GATCACAGCCCTGCACTCCAGCCTGGGTGA---------CAGAGCGAGACTCCATCTCAA 2936

Qy         324 AAACAACAACAAAAAACAAAAAAA 347
               ||  || ||||||||||  | ||
Db        2937 AACGAAAAACAAAAAACAGCACAA 2960


RESULT 46
US-09-949-016-12926/c
; Sequence 12926, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12926
;   LENGTH: 26967
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12926

  Query Match             36.9%;  Score 147.4;  DB 3;  Length 26967;
  Best Local Similarity   72.4%;  Pred. No. 2.4e-33;
  Matches  215;  Conservative    1;  Mismatches   57;  Indels   24;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||
Db       17200 ATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGTGGATCACCTGAGGTCAAGAGT 17141

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || |||||| |||||  |||||||||||||||| ||| ||||||||||||||||| ||||
Db       17140 TCAAGACCAGCCTGGGAAACATGGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 17081

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||| ||||||| | |||||||| |||||||||||| |||| | ||| | | ||
Db       17080 TGGGCGTGGTGGCGCGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGTAGGAGCATCA 17021

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGA---------------------- 297
                 || | |||   | |  |    ||  || |  |||
Db       17020 CTTGAACCCAGGAGGTGGAGGCTGCAGTGAGCCGAGATCATGCCACTGCACTCCAGCCTG 16961

Qy         298 -GTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAA 353
                ||||  ||||  | ||:||||||||||||| || |||||||||||  | | || ||
Db       16960 GGTGACAAGAGCAAAACTCCGTCTCAAAAAAAAAAACAAAAAACAACGACAACAAAA 16904


RESULT 47
US-09-949-016-15831/c
; Sequence 15831, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15831
;   LENGTH: 35104
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-15831

  Query Match             36.9%;  Score 147.4;  DB 3;  Length 35104;
  Best Local Similarity   74.5%;  Pred. No. 2.6e-33;
  Matches  199;  Conservative    0;  Mismatches   66;  Indels    2;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||| |||||| | |||  || ||||||  ||||||||||||
Db        4645 ATGCCTGTAATCCCAGCACTTTGGGAGGACGAGGCAGGTGGATCA--TGAGGTCAAGAGA 4588

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               |||||||||||||||||||||||||||||||||||||||| |||||||||||||| ||||
Db        4587 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAAGTAGC 4528

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||||||||||   |||||||||||||||||||||| | |||   | ||
Db        4527 TGGGCATGGTGGCACGTGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATCG 4468

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                  || |   |   | | ||    ||  || || ||| || ||      | ||
Db        4467 CTTGAACTGGGGAGGTGACGGTTGCAGTGAGCAGAGACTGCGCCACTGCACACCGGCCTG 4408

Qy         321 CAAAAACAACAACAAAAAACAAAAAAA 347
               | || | | ||| | |||| || ||||
Db        4407 CCAACAGAGCAAGACAAAAAAACAAAA 4381


RESULT 48
US-09-949-016-16792
; Sequence 16792, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16792
;   LENGTH: 55114
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16792
```

```
  Query Match            36.9%;  Score 147.4;  DB 3;  Length 55114;
  Best Local Similarity   68.4%;  Pred. No. 3.1e-33;
  Matches  225;  Conservative    1;  Mismatches   87;  Indels   16;  Gaps    1;

Qy         43 TAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTC 102
              || ||| | |   ||  |   | || |     |    | ||||||||||||||| ||||
Db      51385 TACTAAAAGAACATCCAGTTGGCCAGGCGCAGTGGCTCACGCCTGTAATCCCAGTACTTT 51444

Qy        103 GGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACAT 162
              |||||||| |||  |||||||||||||||||||| ||| || |||||| ||||||||||||
Db      51445 GGGAGGCCGAGGCAGGCGGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACAT 51504

Qy        163 GGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTG 222
              |||||||||||||| |||||||||||||||| |||| |||| ||||||||  ||||||
Db      51505 GGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCGGGCGTGGTGGCAGGCACCTG 51564

Qy        223 TAGTCCCAGCTACTCAGGAGCCGGA----------------GATTGCAGTGAGCTGAGAT 266
              || |||||||||||||||| | ||                ||   |||   || ||| |
Db      51565 TAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATGGCTTGAACCCAGGAGGCAGAGGT 51624

Qy        267 CGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAA 326
              |||| | ||   ||  ||||   |  | | | ||  ||||  ||||: || |  |||||
Db      51625 TGCAGTGAGACGAGATCTCACCATTGCCTGGGGGACAAGAGCAAGACTTCGCCAAAAAAA 51684

Qy        327 CAACAACAAAAAACAAAAAAACCATAAGA 355
              || |  |||| | | |||    | || |
Db      51685 AAAAATTAAAATAAATAAATGAAAGAACA 51713


RESULT 49
US-09-949-016-17565
; Sequence 17565, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17565
;   LENGTH: 116955
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17565

  Query Match            36.9%;  Score 147.4;  DB 3;  Length 116955;
  Best Local Similarity   73.1%;  Pred. No. 4.3e-33;
  Matches  201;  Conservative    1;  Mismatches   72;  Indels    1;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| ||||||||||  ||  |||||| |||||| | |||
Db      76957 ACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGACGGTTGGATCACTTGAGGTTAGGAGT 77016

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              |||||||||  |||||| | |||||||||||||   |||| ||||||||||||||||| ||||
Db      77017 TCGAGACCAGCCTGGCTACCATGGTGAAACCTTGTCTCTACTAAAAATACAAAAATTAGC 77076

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||  ||||||  |||||||||| ||||||||||||||||| | |||   | ||
Db      77077 TGGGTGTGGTGGTGCACACCTGTAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAACTG 77136

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCA-CAGAGTGAGCAGAGTGAGACKCCGTC 319
              || | |||   | |  | | | |||||| |    |  | ||||| |||||:|| ||
Db      77137 CTTGAACCCAGGAGGTGGAGGGAGCCGAGATCAGCCCGGACAACAGAGCGAGACTCCATC 77196

Qy        320 TCAAAAACAACAACAAAAAACAAAAAAACCATAAG 354
              ||||| | ||  | | | | | ||    |   ||||
Db      77197 TCAAAGAAAAGGAAATATATCTAAGGCATGGTAAG 77231


RESULT 50
US-09-949-016-13969/c
; Sequence 13969, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13969
;   LENGTH: 301828
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(301828)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13969

  Query Match            36.8%;  Score 147.2;  DB 3;  Length 301828;
  Best Local Similarity   72.6%;  Pred. No. 7.2e-33;
  Matches  209;  Conservative    1;  Mismatches   64;  Indels   14;  Gaps    1;

Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
              |||||||||||||||||| |||||||||| |||||| ||||||||||||||| || |||
Db      69552 CCTGTAATCCCAGCACTTTGGGAGGCCAAAGTGGGCAGATCACCTGAGGTCAAAAGTTCG 69493

Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
              |||||| | |||||||||| |||||||||| |||| |||||||||||||||| ||||  |
Db      69492 AGACCAGCGTGGCCAACACGGTGAAACCCTGTCTCTACTAAAAATACAAAAAGTAGCCAG 69433

Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
              |  ||||||||  |||||||| |||||||||||| |||| | |||   | ||   |
Db      69432 GTGTGGTGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATGGCTT 69373

Qy        264 GATCGCAG--------------AGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTG 309
              || | | |              |||||| || |||||      |   |  | |||||
Db      69372 GAACCCTGGAGACGGAGGTTGCAGTGAGCTGAGATCACGCCCAGCCTGGACAACAGAGCA 69313

Qy        310 AGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
              ||| :|  |||||||| || || |||||| |||||||  | |  | |
Db      69312 AGAGTCTATCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAGAAAA 69265


Search completed: July 17, 2006, 20:41:42
Job time : 170 secs
```

GenCore version 5.1.9

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 21:02:27 ; Search time 1457 Seconds
(without alignments)
3373.407 Million cell updates/sec

Title: SEQ1-47502C
Perfect score: 399.6
Sequence: 1 ccaggtactcagccatgtgc.....tattgcaggagaccaaaaag 400

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : Published_Applications_NA_Main:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	397.2	99.4	36805	8	US-10-741-601-5648	Sequence 5648, Ap
c	2	397.2	99.4	36805	9	US-10-741-600-17657	Sequence 17657, A
c	3	397.2	99.4	75729	8	US-10-741-601-5649	Sequence 5649, Ap
c	4	397.2	99.4	75729	9	US-10-741-600-17658	Sequence 17658, A
c	5	199	49.8	201	8	US-10-741-601-13924	Sequence 13924, A
c	6	199	49.8	201	8	US-10-741-601-14097	Sequence 14097, A
c	7	199	49.8	201	9	US-10-741-600-35944	Sequence 35944, A
c	8	199	49.8	201	9	US-10-741-600-36121	Sequence 36121, A
c	9	197.2	49.3	201	8	US-10-741-601-13993	Sequence 13993, A
c	10	197.2	49.3	201	8	US-10-741-601-14060	Sequence 14060, A
c	11	197.2	49.3	201	9	US-10-741-600-36017	Sequence 36017, A
c	12	197.2	49.3	201	9	US-10-741-600-36084	Sequence 36084, A
c	13	185	46.3	201	8	US-10-741-601-13992	Sequence 13992, A
c	14	185	46.3	201	8	US-10-741-601-14059	Sequence 14059, A
c	15	185	46.3	201	9	US-10-741-600-36016	Sequence 36016, A
c	16	185	46.3	201	9	US-10-741-600-36083	Sequence 36083, A
	17	158.4	39.6	16163	3	US-09-764-891-6505	Sequence 6505, Ap
	18	158.4	39.6	16163	7	US-10-091-414-270	Sequence 270, App
	19	158.4	39.6	96593	3	US-09-997-722-16	Sequence 16, Appl
c	20	158.2	39.6	997	12	US-10-301-480-603905	Sequence 603905,
c	21	158.2	39.6	997	12	US-10-301-480-1217314	Sequence 1217314,
	22	157.4	39.4	635	12	US-10-301-480-577516	Sequence 577516,
	23	157.4	39.4	635	12	US-10-301-480-1190925	Sequence 1190925,
c	24	155.4	38.9	1963	6	US-10-027-632-98771	Sequence 98771, A
c	25	155.4	38.9	1963	7	US-10-027-632-98771	Sequence 98771, A
c	26	155.4	38.9	31898	6	US-10-087-192-760	Sequence 760, App
c	27	154.8	38.7	614	4	US-09-925-065A-532789	Sequence 532789,
c	28	154.8	38.7	614	5	US-09-925-065A-532789	Sequence 532789,
c	29	153.6	38.4	977	12	US-10-301-480-609460	Sequence 609460,
c	30	153.6	38.4	977	12	US-10-301-480-1222869	Sequence 1222869,
c	31	153.4	38.4	415117	10	US-10-995-561-13274	Sequence 13274, A
	32	153	38.3	984	12	US-10-301-480-598906	Sequence 598906,
	33	153	38.3	984	12	US-10-301-480-1212315	Sequence 1212315,
c	34	152.4	38.1	4388	3	US-09-764-891-7647	Sequence 7647, Ap
c	35	152	38.0	174424	3	US-09-967-768A-314	Sequence 314, App
c	36	152	38.0	174424	3	US-09-960-706-969	Sequence 969, App
c	37	152	38.0	174424	10	US-10-843-641A-6459	Sequence 6459, Ap
c	38	152	38.0	181343	9	US-10-723-860-2392	Sequence 2392, Ap
c	39	152	38.0	181343	10	US-10-756-149-2215	Sequence 2215, Ap
	40	151.6	37.9	599	4	US-09-925-065A-45529	Sequence 45529, A
	41	151.6	37.9	599	5	US-09-925-065A-45529	Sequence 45529, A
	42	151.6	37.9	599	12	US-10-301-480-146767	Sequence 146767,
	43	151.6	37.9	599	12	US-10-301-480-760176	Sequence 760176,
c	44	151.4	37.9	169739	8	US-10-450-826-93	Sequence 93, Appl
	45	151	37.8	133300	9	US-10-331-053-70	Sequence 70, Appl
c	46	151	37.8	195998	10	US-10-995-561-13489	Sequence 13489, A
c	47	150.6	37.7	32146	3	US-09-764-860-797	Sequence 797, App
c	48	150.6	37.7	32146	6	US-10-074-095-797	Sequence 797, App
c	49	150.6	37.7	32146	7	US-10-212-872-797	Sequence 797, App
c	50	150.4	37.6	17538	3	US-09-893-348-9	Sequence 9, Appli
c	51	150.4	37.6	17538	3	US-09-218-277-9	Sequence 9, Appli
c	52	150.4	37.6	17538	9	US-10-810-653-9	Sequence 9, Appli
c	53	150.4	37.6	17538	10	US-10-756-149-1809	Sequence 1809, Ap
c	54	150.4	37.6	88892	11	US-10-330-773-672	Sequence 672, App
c	55	150.2	37.6	710	6	US-10-027-632-100019	Sequence 100019,
c	56	150.2	37.6	710	7	US-10-027-632-100019	Sequence 100019,
c	57	150.2	37.6	733	6	US-10-027-632-145872	Sequence 145872,
c	58	150.2	37.6	733	7	US-10-027-632-145872	Sequence 145872,
c	59	149.8	37.5	733	6	US-10-027-632-15604	Sequence 15604, A
c	60	149.8	37.5	733	6	US-10-027-632-15605	Sequence 15605, A
c	61	149.8	37.5	733	6	US-10-027-632-15606	Sequence 15606, A
c	62	149.8	37.5	733	7	US-10-027-632-15604	Sequence 15604, A
c	63	149.8	37.5	733	7	US-10-027-632-15605	Sequence 15605, A
c	64	149.8	37.5	733	7	US-10-027-632-15606	Sequence 15606, A
c	65	149.8	37.5	774	6	US-10-027-632-30004	Sequence 30004, A
c	66	149.8	37.5	774	6	US-10-027-632-30005	Sequence 30005, A
c	67	149.8	37.5	774	7	US-10-027-632-30004	Sequence 30004, A
c	68	149.8	37.5	774	7	US-10-027-632-30005	Sequence 30005, A
	69	149.8	37.5	187844	9	US-10-719-993-6883	Sequence 6883, Ap
	70	149.8	37.5	196686	6	US-10-087-192-484	Sequence 484, App
	71	149.6	37.4	502	6	US-10-027-632-324292	Sequence 324292,
	72	149.6	37.4	502	7	US-10-027-632-324292	Sequence 324292,
c	73	149.6	37.4	5670	3	US-09-954-456-1146	Sequence 1146, Ap
c	74	149.6	37.4	5670	9	US-10-278-698-286	Sequence 286, App
c	75	149.6	37.4	5670	9	US-10-278-698-800	Sequence 800, App
c	76	149.6	37.4	5670	10	US-10-843-641A-4173	Sequence 4173, Ap
c	77	149.6	37.4	50552	9	US-10-719-993-6905	Sequence 6905, Ap
	78	149.6	37.4	493999	9	US-10-719-993-6787	Sequence 6787, Ap
	79	149.2	37.3	502	6	US-10-027-632-6481	Sequence 6481, Ap
	80	149.2	37.3	502	7	US-10-027-632-6481	Sequence 6481, Ap
c	81	149	37.3	8205	3	US-09-860-670-276	Sequence 276, App
c	82	149	37.3	8205	7	US-10-227-646-276	Sequence 276, App
	83	149	37.3	89900	10	US-10-722-939-3	Sequence 3, Appli
	84	149	37.3	285300	10	US-10-857-780-6	Sequence 6, Appli
c	85	148.8	37.2	201143	8	US-10-240-425-1099	Sequence 1099, Ap
c	86	148.6	37.2	44075	3	US-09-997-722-40	Sequence 40, Appl
	87	148.6	37.2	49745	6	US-10-087-192-1930	Sequence 1930, Ap
	88	148.6	37.2	133893	6	US-10-161-510-1	Sequence 1, Appli
c	89	148.2	37.1	965	4	US-09-925-065A-38904	Sequence 38904, A
c	90	148.2	37.1	965	5	US-09-925-065A-38904	Sequence 38904, A

c	91	148.2	37.1	965	12	US-10-301-480-140142 Sequence 140142,
c	92	148.2	37.1	965	12	US-10-301-480-753551 Sequence 753551,
	93	148.2	37.1	206739	10	US-10-981-277-49 Sequence 49, Appl
	94	148	37.0	143973	6	US-10-087-192-442 Sequence 442, App
	95	147.8	37.0	539	12	US-10-301-480-221859 Sequence 221859,
	96	147.8	37.0	539	12	US-10-301-480-835268 Sequence 835268,
	97	147.8	37.0	568	4	US-09-925-065A-124093 Sequence 124093,
	98	147.8	37.0	568	5	US-09-925-065A-124093 Sequence 124093,
	99	147.8	37.0	177900	10	US-10-287-436A-1277 Sequence 1277, Ap
	100	147.6	36.9	709	10	US-10-956-157-3291 Sequence 3291, Ap
	101	147.6	36.9	709	10	US-10-956-157-8526 Sequence 8526, Ap
c	102	147.6	36.9	983	12	US-10-301-480-538078 Sequence 538078,
c	103	147.6	36.9	983	12	US-10-301-480-1151487 Sequence 1151487,
	104	147.6	36.9	11820	10	US-10-756-149-2006 Sequence 2006, Ap
	105	147.6	36.9	78268	6	US-10-087-192-742 Sequence 742, App
	106	147.6	36.9	180531	16	US-11-114-798-57 Sequence 57, Appl
c	107	147.4	36.9	1002	16	US-11-043-752-3751 Sequence 3751, Ap
	108	147.4	36.9	90798	8	US-10-318-819A-4 Sequence 4, Appli
c	109	147.4	36.9	98634	6	US-10-087-192-1870 Sequence 1870, Ap
	110	147.2	36.8	47219	8	US-10-699-941-10 Sequence 10, Appl
	111	147.2	36.8	59884	8	US-10-699-941-11 Sequence 11, Appl
c	112	147.2	36.8	599001	10	US-10-317-869A-4 Sequence 4, Appli
c	113	147.2	36.8	775062	9	US-10-719-993-6844 Sequence 6844, Ap
c	114	147.2	36.8	783062	10	US-10-461-862-166 Sequence 166, App
	115	147	36.8	619	12	US-10-301-480-580582 Sequence 580582,
	116	147	36.8	619	12	US-10-301-480-1193991 Sequence 1193991,
	117	147	36.8	622	4	US-09-925-065A-741359 Sequence 741359,
	118	147	36.8	622	5	US-09-925-065A-741359 Sequence 741359,
c	119	147	36.8	965	4	US-09-925-065A-38903 Sequence 38903, A
c	120	147	36.8	965	5	US-09-925-065A-38903 Sequence 38903, A
c	121	147	36.8	965	12	US-10-301-480-140141 Sequence 140141,
c	122	147	36.8	965	12	US-10-301-480-753550 Sequence 753550,
c	123	147	36.8	997	4	US-09-925-065A-93228 Sequence 93228, A
c	124	147	36.8	997	4	US-09-925-065A-93229 Sequence 93229, A
c	125	147	36.8	997	4	US-09-925-065A-93230 Sequence 93230, A
c	126	147	36.8	997	5	US-09-925-065A-93228 Sequence 93228, A
c	127	147	36.8	997	5	US-09-925-065A-93229 Sequence 93229, A
c	128	147	36.8	997	5	US-09-925-065A-93230 Sequence 93230, A
c	129	147	36.8	997	12	US-10-301-480-194470 Sequence 194470,
c	130	147	36.8	997	12	US-10-301-480-194471 Sequence 194471,
c	131	147	36.8	997	12	US-10-301-480-194472 Sequence 194472,
c	132	147	36.8	997	12	US-10-301-480-807879 Sequence 807879,
c	133	147	36.8	997	12	US-10-301-480-807880 Sequence 807880,
c	134	147	36.8	997	12	US-10-301-480-807881 Sequence 807881,
c	135	147	36.8	2909	7	US-10-108-260A-1113 Sequence 1113, Ap
c	136	147	36.8	45664	9	US-10-741-600-17881 Sequence 17881, A
c	137	147	36.8	57013	6	US-10-087-192-1798 Sequence 1798, Ap
	138	147	36.8	238417	10	US-10-461-862-98 Sequence 98, Appl
c	139	146.8	36.7	442	9	US-10-674-124A-4513 Sequence 4513, Ap
c	140	146.8	36.7	505	6	US-10-027-632-229704 Sequence 229704,
c	141	146.8	36.7	505	7	US-10-027-632-229704 Sequence 229704,
c	142	146.8	36.7	613	4	US-09-925-065A-634598 Sequence 634598,
c	143	146.8	36.7	613	5	US-09-925-065A-634598 Sequence 634598,
	144	146.8	36.7	2061	3	US-09-984-245-14 Sequence 14, Appl
	145	146.8	36.7	2061	3	US-09-966-262-14 Sequence 14, Appl
	146	146.8	36.7	2061	3	US-09-983-966-14 Sequence 14, Appl
	147	146.8	36.7	2061	6	US-10-059-395-14 Sequence 14, Appl
	148	146.8	36.7	2061	6	US-10-143-090-14 Sequence 14, Appl
	149	146.8	36.7	2061	10	US-10-960-251-14 Sequence 14, Appl
c	150	146.6	36.7	515	4	US-09-925-065A-853585 Sequence 853585,

ALIGNMENTS

```
RESULT 1
US-10-741-601-5648/c
; Sequence 5648, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5648
;   LENGTH: 36805
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(36805)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-601-5648

  Query Match             99.4%;  Score 397.2;  DB 8;  Length 36805;
  Best Local Similarity   99.2%;  Pred. No. 2.2e-111;
  Matches  397;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCAGGTACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27800 CCAGGTACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGG 27741

Qy         61 CCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCG 120
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      27740 CCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCG 27681

Qy        121 GATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTA 180
              ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27680 GATCACSTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTA 27621

Qy        181 CTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27620 CTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGG 27561

Qy        241 AGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27560 AGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTG 27501

Qy        301 AGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 360
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27500 AGSAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 27441

Qy        361 TCCATCTGCGGTTCCCAGACTATTGCAGGAGACCAAAAAG 400
              |||||||||||||||||||||||||||||||||||||||
Db      27440 TCCATCTGCGGTTCCCAGACTATTGCAGGAGACCAAAAAG 27401


RESULT 2
US-10-741-600-17657/c
; Sequence 17657, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17657
;   LENGTH: 36805
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(36805)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17657

  Query Match             99.4%;  Score 397.2;  DB 9;  Length 36805;
  Best Local Similarity   99.2%;  Pred. No. 2.2e-111;
```

```
  Matches  397;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCAGGTACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27800 CCAGGTACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGG 27741

Qy         61 CCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCG 120
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      27740 CCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCG 27681

Qy        121 GATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTA 180
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27680 GATCACSTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTA 27621

Qy        181 CTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27620 CTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGG 27561

Qy        241 AGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27560 AGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTG 27501

Qy        301 AGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 360
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27500 AGSAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 27441

Qy        361 TCCATCTGCGGTTCCCAGACTATTGCAGGAGACCAAAAAG 400
              |||||||||||||||||||||||||||||||||||||||
Db      27440 TCCATCTGCGGTTCCCAGACTATTGCAGGAGACCAAAAAG 27401


RESULT 3
US-10-741-601-5649/c
; Sequence 5649, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5649
;   LENGTH: 75729
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(75729)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-601-5649

  Query Match             99.4%;  Score 397.2;  DB 8;  Length 75729;
  Best Local Similarity   99.2%;  Pred. No. 3e-111;
  Matches  397;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCAGGTACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      59242 CCAGGTACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGG 59183

Qy         61 CCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCG 120
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      59182 CCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCG 59123

Qy        121 GATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTA 180
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      59122 GATCACSTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTA 59063

Qy        181 CTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      59062 CTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGG 59003

Qy        241 AGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      59002 AGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTG 58943

Qy        301 AGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 360
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58942 AGSAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 58883

Qy        361 TCCATCTGCGGTTCCCAGACTATTGCAGGAGACCAAAAAG 400
              |||||||||||||||||||||||||||||||||||||||
Db      58882 TCCATCTGCGGTTCCCAGACTATTGCAGGAGACCAAAAAG 58843


RESULT 4
US-10-741-600-17658/c
; Sequence 17658, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17658
;   LENGTH: 75729
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(75729)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17658

  Query Match             99.4%;  Score 397.2;  DB 9;  Length 75729;
  Best Local Similarity   99.2%;  Pred. No. 3e-111;
  Matches  397;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCAGGTACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      59242 CCAGGTACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGG 59183

Qy         61 CCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCG 120
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      59182 CCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCG 59123

Qy        121 GATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTA 180
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      59122 GATCACSTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTA 59063

Qy        181 CTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      59062 CTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGG 59003

Qy        241 AGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      59002 AGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTG 58943

Qy        301 AGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 360
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58942 AGSAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 58883

Qy        361 TCCATCTGCGGTTCCCAGACTATTGCAGGAGACCAAAAAG 400
              |||||||||||||||||||||||||||||||||||||||
Db      58882 TCCATCTGCGGTTCCCAGACTATTGCAGGAGACCAAAAAG 58843


RESULT 5
US-10-741-601-13924/c
; Sequence 13924, Application US/10741601
```

```
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13924
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-601-13924

  Query Match             49.8%;  Score 199;  DB 8;  Length 201;
  Best Local Similarity   99.0%;  Pred. No. 5.9e-51;
  Matches  199;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         27 ATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCT 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCT 142

Qy         87 GTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGA 146
              |||||||||||||| |||||||||||||||||||:||||||||||||||||||||||||
Db        141 GTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCGGATCACSTGAGGTCAAGAGATCGAGA 82

Qy        147 CCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCA 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 CCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCA 22

Qy        207 TGGTGGCACACACCTGTAGTC 227
              |||||||||||||||||||||
Db         21 TGGTGGCACACACCTGTAGTC 1


RESULT 6
US-10-741-601-14097/c
; Sequence 14097, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14097
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-601-14097

  Query Match             49.8%;  Score 199;  DB 8;  Length 201;
  Best Local Similarity   99.0%;  Pred. No. 5.9e-51;
  Matches  199;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         27 ATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCT 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCT 142

Qy         87 GTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGA 146
              |||||||||||||| |||||||||||||||||||:||||||||||||||||||||||||
Db        141 GTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCGGATCACSTGAGGTCAAGAGATCGAGA 82

Qy        147 CCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCA 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 CCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCA 22

Qy        207 TGGTGGCACACACCTGTAGTC 227
              |||||||||||||||||||||
Db         21 TGGTGGCACACACCTGTAGTC 1


RESULT 7
US-10-741-600-35944/c
; Sequence 35944, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 35944
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-35944

  Query Match             49.8%;  Score 199;  DB 9;  Length 201;
  Best Local Similarity   99.0%;  Pred. No. 5.9e-51;
  Matches  199;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         27 ATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCT 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCT 142

Qy         87 GTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGA 146
              |||||||||||||| |||||||||||||||||||:||||||||||||||||||||||||
Db        141 GTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCGGATCACSTGAGGTCAAGAGATCGAGA 82

Qy        147 CCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCA 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 CCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCA 22

Qy        207 TGGTGGCACACACCTGTAGTC 227
              |||||||||||||||||||||
Db         21 TGGTGGCACACACCTGTAGTC 1


RESULT 8
US-10-741-600-36121/c
; Sequence 36121, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36121
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-36121

  Query Match             49.8%;  Score 199;  DB 9;  Length 201;
  Best Local Similarity   99.0%;  Pred. No. 5.9e-51;
  Matches  199;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         27 ATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCT 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCT 142

Qy         87 GTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGA 146
              |||||||||||||| |||||||||||||||||||:||||||||||||||||||||||||
```

```
Db        141 GTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCGGATCACSTGAGGTCAAGAGATCGAGA 82

Qy        147 CCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCA 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 CCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCA 22

Qy        207 TGGTGGCACACACCTGTAGTC 227
              |||||||||||||||||||||
Db         21 TGGTGGCACACACCTGTAGTC 1


RESULT 9
US-10-741-601-13993/c
; Sequence 13993, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13993
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-601-13993

  Query Match           49.3%;  Score 197.2;  DB 8;  Length 201;
  Best Local Similarity   99.0%;  Pred. No. 2.1e-50;
  Matches  196;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 142

Qy        263 AGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCA 322
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||:||||||||
Db        141 AGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGSAGAGTGAGACTCCGTCTCA 82

Qy        323 AAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTA 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTA 22

Qy        383 TTGCAGGAGACCAAAAAG 400
              ||||||||||||||||||
Db         21 TTGCAGGAGACCAAAAAG 4


RESULT 10
US-10-741-601-14060/c
; Sequence 14060, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14060
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-601-14060

  Query Match           49.3%;  Score 197.2;  DB 8;  Length 201;
  Best Local Similarity   99.0%;  Pred. No. 2.1e-50;
  Matches  196;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 142

Qy        263 AGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCA 322
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||:||||||||
Db        141 AGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGSAGAGTGAGACTCCGTCTCA 82

Qy        323 AAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTA 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTA 22

Qy        383 TTGCAGGAGACCAAAAAG 400
              ||||||||||||||||||
Db         21 TTGCAGGAGACCAAAAAG 4


RESULT 11
US-10-741-600-36017/c
; Sequence 36017, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36017
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-36017

  Query Match           49.3%;  Score 197.2;  DB 9;  Length 201;
  Best Local Similarity   99.0%;  Pred. No. 2.1e-50;
  Matches  196;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 142

Qy        263 AGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCA 322
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||:||||||||
Db        141 AGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGSAGAGTGAGACTCCGTCTCA 82

Qy        323 AAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTA 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTA 22

Qy        383 TTGCAGGAGACCAAAAAG 400
              ||||||||||||||||||
Db         21 TTGCAGGAGACCAAAAAG 4


RESULT 12
US-10-741-600-36084/c
; Sequence 36084, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36084
;   LENGTH: 201
```

```
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-10-741-600-36084

  Query Match             49.3%;  Score 197.2;  DB 9;  Length 201;
  Best Local Similarity   99.0%;  Pred. No. 2.1e-50;
  Matches  196;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTG 142

Qy        263 AGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCA 322
              ||||||||||||||||||||||||||||||||||||||:|||||||||||:||||||||
Db        141 AGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGSAGAGTGAGACTCCGTCTCA 82

Qy        323 AAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTA 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTA 22

Qy        383 TTGCAGGAGACCAAAAAG 400
              |||||||||||||||||
Db         21 TTGCAGGAGACCAAAAAG 4


RESULT 13
US-10-741-601-13992/c
; Sequence 13992, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13992
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-601-13992

  Query Match             46.3%;  Score 185;  DB 8;  Length 201;
  Best Local Similarity   99.5%;  Pred. No. 1.3e-46;
  Matches  186;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        214 ACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAG 273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAG 142

Qy        274 TGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAAC 333
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        141 TGAGCCGAAATCACAGATCACAGAGTGAGGAGAGTGAGACKCCGTCTCAAAAACAACAAC 82

Qy        334 AAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTATTGCAGGAGAC 393
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTATTGCAGGAGAC 22

Qy        394 CAAAAAG 400
              |||||||
Db         21 CAAAAAG 15


RESULT 14
US-10-741-601-14059/c
; Sequence 14059, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14059
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-601-14059

  Query Match             46.3%;  Score 185;  DB 8;  Length 201;
  Best Local Similarity   99.5%;  Pred. No. 1.3e-46;
  Matches  186;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        214 ACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAG 273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAG 142

Qy        274 TGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAAC 333
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        141 TGAGCCGAAATCACAGATCACAGAGTGAGGAGAGTGAGACKCCGTCTCAAAAACAACAAC 82

Qy        334 AAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTATTGCAGGAGAC 393
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTATTGCAGGAGAC 22

Qy        394 CAAAAAG 400
              |||||||
Db         21 CAAAAAG 15


RESULT 15
US-10-741-600-36016/c
; Sequence 36016, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36016
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-36016

  Query Match             46.3%;  Score 185;  DB 9;  Length 201;
  Best Local Similarity   99.5%;  Pred. No. 1.3e-46;
  Matches  186;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        214 ACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAG 273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAG 142

Qy        274 TGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAAC 333
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        141 TGAGCCGAAATCACAGATCACAGAGTGAGGAGAGTGAGACKCCGTCTCAAAAACAACAAC 82

Qy        334 AAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTATTGCAGGAGAC 393
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTATTGCAGGAGAC 22

Qy        394 CAAAAAG 400
              |||||||
Db         21 CAAAAAG 15


RESULT 16
```

```
US-10-741-600-36083/c
; Sequence 36083, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36083
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-36083

  Query Match          46.3%;  Score 185;  DB 9;  Length 201;
  Best Local Similarity   99.5%;  Pred. No. 1.3e-46;
  Matches  186;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        214 ACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAG 273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAG 142

Qy        274 TGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAAC 333
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        141 TGAGCCGAAATCACAGATCACAGAGTGAGGAGAGTGAGACKCCGTCTCAAAAACAACAAC 82

Qy        334 AAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTATTGCAGGAGAC 393
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGACTATTGCAGGAGAC 22

Qy        394 CAAAAAG 400
              |||||||
Db         21 CAAAAAG 15


RESULT 17
US-09-764-891-6505
; Sequence 6505, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6505
;   LENGTH: 16163.
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-891-6505

  Query Match          39.6%;  Score 158.4;  DB 3;  Length 16163;
  Best Local Similarity   73.9%;  Pred. No. 1.2e-37;
  Matches  201;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy         76 GCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCA 135
              | |  | ||||||||||||||||||  |||||||| ||| ||| |||||||||||| ||||
Db       8172 GGTTCACGCCTGTAATCCCAGCACTCTGGGAGGCCGAGGCGGGTGGATCACCTGATGTCA 8231

Qy        136 AGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAA 195
              ||| | |||| ||  |||||||||||||||||||||||||||| ||||||| |||||||||
Db       8232 GGAGTTTGAGATCAGTCTGGCCAACATGGTGAAACCCCGTCTCTACTAAATATACAAAAA 8291

Qy        196 ATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG 255
              |||||||  |||||| |||| | |||| ||||||||||| |||| ||||| ||||||
Db       8292 CCAGCTGGGTGTGGTGGTACACGCTTGTAATCCCAGCTACTCTGGAGGCGGAGGTTGCAG 8351

Qy        256 TGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKC 315
              ||||| || |  |||       |  |  | |   | | ||||||| |||          |
Db       8352 TGAGCCGAAACTGCACCACTACACTCCAGCCTGGGTGACAGAGGGAGACTCTGTCTCCCA 8411

Qy        316 CGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
              |    ||||| || || ||||||| |||||||
Db       8412 GGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 8443


RESULT 18
US-10-091-414-270
; Sequence 270, Application US/10091414
; Publication No. US20030224461A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA116C1
;  CURRENT APPLICATION NUMBER: US/10/091,414
;  CURRENT FILING DATE:  2002-03-07
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 392
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 270
;   LENGTH: 16163
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-091-414-270

  Query Match          39.6%;  Score 158.4;  DB 7;  Length 16163;
  Best Local Similarity   73.9%;  Pred. No. 1.2e-37;
  Matches  201;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy         76 GCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCA 135
              | |  | ||||||||||||||||||  |||||||| ||| ||| |||||||||||| ||||
Db       8172 GGTTCACGCCTGTAATCCCAGCACTCTGGGAGGCCGAGGCGGGTGGATCACCTGATGTCA 8231

Qy        136 AGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAA 195
              ||| | |||| ||  |||||||||||||||||||||||||||| ||||||| |||||||||
Db       8232 GGAGTTTGAGATCAGTCTGGCCAACATGGTGAAACCCCGTCTCTACTAAATATACAAAAA 8291

Qy        196 ATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG 255
              |||||||  |||||| |||| | |||| ||||||||||| |||| ||||| ||||||
Db       8292 CCAGCTGGGTGTGGTGGTACACGCTTGTAATCCCAGCTACTCTGGAGGCGGAGGTTGCAG 8351

Qy        256 TGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKC 315
              ||||| || |  |||       |  |  | |   | | ||||||| |||          |
Db       8352 TGAGCCGAAACTGCACCACTACACTCCAGCCTGGGTGACAGAGGGAGACTCTGTCTCCCA 8411

Qy        316 CGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
              |    ||||| || || ||||||| |||||||
Db       8412 GGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 8443


RESULT 19
US-09-997-722-16
; Sequence 16, Application US/09997722
; Publication No. US20040072154A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David
;  APPLICANT:  Engelhard, Eric
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
;  FILE REFERENCE: A-71171/RMS/DCF
;  CURRENT APPLICATION NUMBER: US/09/997,722
;  CURRENT FILING DATE:  2001-11-30
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 301
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
;   LENGTH: 96593
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-997-722-16

  Query Match             39.6%;  Score 158.4;  DB 3;  Length 96593;
  Best Local Similarity   75.0%;  Pred. No. 2.5e-37;
  Matches  216;  Conservative    1;  Mismatches   57;  Indels   14;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||||||||||||||||| |||||||| ||| ||||||||  |||||||||| |||
Db       18217 ACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATTTCCTGAGGTCAGGAGT 18276

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               || ||| || |||||||||||||||||||||||||||| ||||||||||||||||| ||||
Db       18277 TCAAGATCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 18336

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||||||||   |||||||||||||||||||| |||| | |||   | ||
Db       18337 TGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCG 18396

Qy         261 TGAGATCGCAG--------------AGTGAGCCGAAATCACAGATCACAGAGTGAGCAGA 306
                  || | |||              |||||||||| ||| ||        | ||||  ||
Db       18397 CTTGAACCCAGGCGGCGGAGGTTGCAGTGAGCCGAGATCGCACCAGCTTGGGTGACAAGC 18456

Qy         307 GTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAG 354
               |||||||: | ||||||||| ||   | |||| | |||| ||||| | |
Db       18457 GTGAGACTTCATCTCAAAAAAAAACAAAAAACAAAAAACAACCAGATG 18504


RESULT 20
US-10-301-480-603905/c
; Sequence 603905, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 603905
;   LENGTH: 997
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-603905

  Query Match             39.6%;  Score 158.2;  DB 12;  Length 997;
  Best Local Similarity   78.5%;  Pred. No. 4.7e-38;
  Matches  205;  Conservative    0;  Mismatches   48;  Indels    8;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||| |||||||  |||  ||||||| ||||||||||  |||
Db         733 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAGGCGGATTACCTGAGGTCTGGAGT 674

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               || |||||| |||||||||||||||||||||| |||| ||||||||||||||||  |||
Db         673 TCAAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATCAGC 614

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||||||||    |||||| |||||||||||||||| ||||| ||||||||||
Db         613 TGGGCATGGTGGCAGGGGCCTGTAATCCCAGCTACTCAGGAGGCGGAGGTTGCAGTGAGC 554

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               |||||| |          | |   |||      | | ||||  ||||  | || | ||||
Db         553 TGAGATTG--------TGCCACTGCACTCCAGCCTGGGTGACAAGAGCAAAACACTGTCT 502

Qy         321 CAAAAACAACAACAAAAAACA 341
               |||||| || || |||| |||
Db         501 CAAAAAAAAAAAAAAAAGACA 481


RESULT 21
US-10-301-480-1217314/c
; Sequence 1217314, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1217314
;   LENGTH: 997
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1217314

  Query Match             39.6%;  Score 158.2;  DB 12;  Length 997;
  Best Local Similarity   78.5%;  Pred. No. 4.7e-38;
  Matches  205;  Conservative    0;  Mismatches   48;  Indels    8;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||| |||||||  |||  ||||||| ||||||||||  |||
Db         733 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAGGCGGATTACCTGAGGTCTGGAGT 674

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               || |||||| |||||||||||||||||||||| |||| ||||||||||||||||  |||
Db         673 TCAAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATCAGC 614

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |||||||||||||    |||||| |||||||||||||||| ||||| ||||||||||
Db         613 TGGGCATGGTGGCAGGGGCCTGTAATCCCAGCTACTCAGGAGGCGGAGGTTGCAGTGAGC 554

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               |||||| |          | |   |||      | | ||||  ||||  | || | ||||
Db         553 TGAGATTG--------TGCCACTGCACTCCAGCCTGGGTGACAAGAGCAAAACACTGTCT 502

Qy         321 CAAAAACAACAACAAAAAACA 341
               |||||| || || |||| |||
Db         501 CAAAAAAAAAAAAAAAAGACA 481


RESULT 22
US-10-301-480-577516
; Sequence 577516, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 577516
;   LENGTH: 635
;   TYPE: DNA
;   ORGANISM: Homo sapien
```

US-10-301-480-577516

Query Match 39.4%; Score 157.4; DB 12; Length 635;
Best Local Similarity 77.2%; Pred. No. 6.9e-38;
Matches 206; Conservative 1; Mismatches 52; Indels 8; Gaps 1;

```
Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||||||| ||| |||| |||||||||||| |||
Db        252 ACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGTGGATAACCTGAGGTCAGGAGT 311

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || |||||| ||||||||| |||||||||||| ||| ||||||||||||||||| ||||
Db        312 TCAAGACCAGCCTGGCCAAAATGGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 371

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||| ||||    ||||||| |||||||||||| |||| ||||| |||| ||||||
Db        372 TGGGCATGATGGCGGGTACCTGTAATCCCAGCTACTCGGGAGGCGGAGGTTGCGGTGAGC 431

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||||||          | |   |||       | | |  |  |||| || ||:| ||||
Db        432 CGAGATCG--------TGCCATTGCACCCCAGCCTGGGCAACAAGAGCGAAACTCTGTCT 483

Qy        321 CAAAAACAACAACAAAAAACAAAAAAA 347
              |||||| || |  ||| || ||| |||
Db        484 CAAAAAAAAAAGAAAAGAAAAAAGAAA 510
```

RESULT 23
US-10-301-480-1190925
; Sequence 1190925, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1190925
; LENGTH: 635
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1190925

Query Match 39.4%; Score 157.4; DB 12; Length 635;
Best Local Similarity 77.2%; Pred. No. 6.9e-38;
Matches 206; Conservative 1; Mismatches 52; Indels 8; Gaps 1;

```
Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||||||| ||| |||| |||||||||||| |||
Db        252 ACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGTGGATAACCTGAGGTCAGGAGT 311

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || |||||| ||||||||| |||||||||||| ||| ||||||||||||||||| ||||
Db        312 TCAAGACCAGCCTGGCCAAAATGGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 371

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||| ||||    ||||||| |||||||||||| |||| ||||| |||| ||||||
Db        372 TGGGCATGATGGCGGGTACCTGTAATCCCAGCTACTCGGGAGGCGGAGGTTGCGGTGAGC 431

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||||||          | |   |||       | | |  |  |||| || ||:| ||||
Db        432 CGAGATCG--------TGCCATTGCACCCCAGCCTGGGCAACAAGAGCGAAACTCTGTCT 483

Qy        321 CAAAAACAACAACAAAAAACAAAAAAA 347
              |||||| || |  ||| || ||| |||
Db        484 CAAAAAAAAAAGAAAAGAAAAAAGAAA 510
```

RESULT 24
US-10-027-632-98771/c
; Sequence 98771, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 98771
; LENGTH: 1963
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-98771

Query Match 38.9%; Score 155.4; DB 6; Length 1963;
Best Local Similarity 72.6%; Pred. No. 4.5e-37;
Matches 217; Conservative 1; Mismatches 72; Indels 9; Gaps 1;

```
Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||| |||||||||||||| || ||||||||||||||||||| ||||| || ||
Db        463 ACGCCTATAATCCCAGCACTTTGGAAGGCCAAGGTGGGCGGATCACTTGAGGCCAGGAAT 404

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||||||| |||||||||||||||||| ||||||||||||||||| ||||
Db        403 TCGAGACCAGCCTGGCCAAAATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 344

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
                ||  ||||||    | |||||||||||| |||||| |||| ||||| ||||||||||
Db        343 CCGGAGTGGTGGTGGGCGCCTGTAGTCCCAACTACTCGGGAGGCGGAGGTTGCAGTGAGC 284

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              ||||  |  |   |  | |  |  |         |||||  |||||:| ||||
Db        283 CGAGACCACGCCTTTGCACTCCAGCCTGGGCGA---------CAGAGCGAGACTCTGTCT 233

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGA 379
              ||| ||||| ||||||||| ||| ||   |||  ||| | |  |  |  ||||
Db        232 CAACAACAATAACAAAAAAAGAAACAAGTGCAAGGTCTTGCCTTTGAGTTGGGGTCAGA 174
```

RESULT 25
US-10-027-632-98771/c
; Sequence 98771, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12

```
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 98771
;   LENGTH: 1963
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-98771

  Query Match             38.9%;  Score 155.4;  DB 7;  Length 1963;
  Best Local Similarity   72.6%;  Pred. No. 4.5e-37;
  Matches  217;  Conservative    1;  Mismatches   72;  Indels    9;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | ||||  ||||||||||||||| || ||||||||||||||||||||| |||||  || ||
Db        463 ACGCCTATAATCCCAGCACTTTGGAAGGCCAAGGTGGGCGGATCACTTGAGGCCAGGAAT 404

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              |||||||||  |||||||||  |||||||||||||||||  |||||||||||||||||  ||||
Db        403 TCGAGACCAGCCTGGCCAAAATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 344

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
                ||  ||||||    | ||||||||||||  ||||||  ||||  |||||  ||||||||||
Db        343 CCGGAGTGGTGGTGGGCGCCTGTAGTCCCAACTACTCGGGAGGCGGAGGTTGCAGTGAGC 284

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              ||||  | |    |    |   | |   |   |          |||||  |||||:|  ||||
Db        283 CGAGACCACGCCTTTGCACTCCAGCCTGGGCGA---------CAGAGCGAGACTCTGTCT 233

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGA 379
              |||  |||||  |||||||||  |||  ||    |||   |||  |  |  |  |    ||||
Db        232 CAACAACAATAACAAAAAAAGAAACAAGTGCAAGGTCTTGCCTTTGAGTTGGGGTCAGA 174


RESULT 26
US-10-087-192-760/c
; Sequence 760, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 760
;   LENGTH: 31898
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-087-192-760

  Query Match             38.9%;  Score 155.4;  DB 6;  Length 31898;
  Best Local Similarity   72.6%;  Pred. No. 1.4e-36;
  Matches  217;  Conservative    1;  Mismatches   72;  Indels    9;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | ||||  ||||||||||||||| || ||||||||||||||||||||| |||||  || ||
Db      30243 ACGCCTATAATCCCAGCACTTTGGAAGGCCAAGGTGGGCGGATCACTTGAGGCCAGGAAT 30184

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              |||||||||  |||||||||  |||||||||||||||||  |||||||||||||||||  ||||
Db      30183 TCGAGACCAGCCTGGCCAAAATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 30124

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
                ||  ||||||    | ||||||||||||  ||||||  ||||  |||||  ||||||||||
Db      30123 CCGGAGTGGTGGTGGGCGCCTGTAGTCCCAACTACTCGGGAGGCGGAGGTTGCAGTGAGC 30064

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              ||||  | |    |    |   | |   |   |          |||||  |||||:|  ||||
Db      30063 CGAGACCACGCCTTTGCACTCCAGCCTGGGCGA---------CAGAGCGAGACTCTGTCT 30013

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTCCCAGA 379
              |||  |||||  |||||||||  |||  ||    |||   |||  |  |  |  |    ||||
Db      30012 CAACAACAATAACAAAAAAAGAAACAAGTGCAAGGTCTTGCCTTTGAGTTGGGGTCAGA 29954


RESULT 27
US-09-925-065A-532789/c
; Sequence 532789, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 532789
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-532789

  Query Match             38.7%;  Score 154.8;  DB 4;  Length 614;
  Best Local Similarity   70.3%;  Pred. No. 4.3e-37;
  Matches  218;  Conservative    2;  Mismatches   89;  Indels    1;  Gaps    1;

Qy         48 AGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAG 107
              |||  |  ||  |||  | |  |      | || | ||||||||||||||||||  |||||
Db        614 AGAAACAGTGTGGCAGGCCGGGTGCGGTGGTGCACACCTGTAATCCCAGCACTTTGGGAG 555

Qy        108 GCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGA 167
              ||||||||||||||||  ||||||  ||  |||  |||||||||  ||||||||||||||  |
Db        554 GCCAAGGTGGGCGGATCGCCTGAGATCAGGAGTTCGAGACCAGCCTGGCCAACATGGCAA 495

Qy        168 AACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTC 227
              ||||||  |||  ||||||||||||||||  |||||||    | ||||||     |||||| ||
Db        494 AACCCCATCTCTACTAAAAATACAAAAATTAGCTGGATGTTGTGGCAGGTGCCTGTAATC 435

Qy        228 CCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCAC 287
              ||||||||||||||  |||   | ||          || |   |  | |  |    |||
Db        434 CCAGCTACTCAGGAGGCTGAGGCAGGAGAATCTCTTGAACCTGGGAGGCGGAGGTTGCAC 375

Qy        288 AGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
                | | || |||||  ||||:| |||||||||| || || |||||| |:|||||
Db        374 TCCAGCCTGGGAGA-CAGAGGAAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAARAAAAA 316
```

```
Qy        348 CCATAAGACA 357
              | || | |
Db        315 GAAGAAAAAA 306


RESULT 28
US-09-925-065A-532789/c
; Sequence 532789, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 532789
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-532789

  Query Match             38.7%;  Score 154.8;  DB 5;  Length 614;
  Best Local Similarity   70.3%;  Pred. No. 4.3e-37;
  Matches  218;  Conservative    2;  Mismatches   89;  Indels    1;  Gaps    1;

Qy         48 AGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAG 107
              ||| |  ||  |||  | |  |      | || |  ||||||||||||||||||| |||||
Db        614 AGAAACAGTGTGGCAGGCCGGGTGCGGTGGTGCACACCTGTAATCCCAGCACTTTGGGAG 555

Qy        108 GCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGA 167
              |||||||||||||||||| |||||| ||| ||| |||||||| ||||||||||||||  |
Db        554 GCCAAGGTGGGCGGATCGCCTGAGATCAGGAGTTCGAGACCAGCCTGGCCAACATGGCAA 495

Qy        168 AACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTC 227
              |||||| ||| |||||||||||||||||| |||||||    | ||||||     |||||| ||
Db        494 AACCCCATCTCTACTAAAAATACAAAAATTAGCTGGATGTTGTGGCAGGTGCCTGTAATC 435

Qy        228 CCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCAC 287
              |||||||||||||||| | |||    | ||          || |   |   | |   |||
Db        434 CCAGCTACTCAGGAGGCTGAGGCAGGAGAATCTCTTGAACCTGGGAGGCGGAGGTTGCAC 375

Qy        288 AGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
                | | | || |||||  ||||:| |||||||||| || || ||||||  |:|||||
Db        374 TCCAGCCTGGGAGA-CAGAGGAAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAARAAAAA 316

Qy        348 CCATAAGACA 357
              | || | |
Db        315 GAAGAAAAAA 306


RESULT 29
US-10-301-480-609460/c
; Sequence 609460, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 609460
;   LENGTH: 977
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-609460

  Query Match             38.4%;  Score 153.6;  DB 12;  Length 977;
  Best Local Similarity   70.3%;  Pred. No. 1.2e-36;
  Matches  218;  Conservative    1;  Mismatches   90;  Indels    1;  Gaps    1;

Qy         48 AGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAG 107
              ||| |  ||  |||  | |  |      | || |  ||||||||||||||||||| |||||
Db        436 AGAAACAGTGTGGCAGGCCGGGTGCGGTGGTGCACACCTGTAATCCCAGCACTTTGGGAG 377

Qy        108 GCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGA 167
              |||||||||||||||||| |||||| ||| ||| |||||||| ||||||||||||||  |
Db        376 GCCAAGGTGGGCGGATCGCCTGAGATCAGGAGTTCGAGACCAGCCTGGCCAACATGGCAA 317

Qy        168 AACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTC 227
              |||||| ||| |||||||||||||||||| |||||||    | ||||||     |||||| ||
Db        316 AACCCCATCTCTACTAAAAATACAAAAATTAGCTGGATGTTGTGGCAGGTGCCTGTAATC 257

Qy        228 CCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCAC 287
              |||||||||||||||| | |||    | ||          || |   |   | |   |||
Db        256 CCAGCTACTCAGGAGGCTGAGGCAGGAGAATCTCTTGAACCTGGGAGGCGGAGGTTGCAC 197

Qy        288 AGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
                | | | || |||||  ||||:| |||||||||| || || ||||||  |:|||||
Db        196 TCCAGCCTGGGAGA-CAGAGGAAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAARAAAAA 138

Qy        348 CCATAAGACA 357
              | || | |
Db        137 GAAGAAAAAA 128


RESULT 30
US-10-301-480-1222869/c
; Sequence 1222869, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1222869
;   LENGTH: 977
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1222869

  Query Match             38.4%;  Score 153.6;  DB 12;  Length 977;
  Best Local Similarity   70.3%;  Pred. No. 1.2e-36;
  Matches  218;  Conservative    1;  Mismatches   90;  Indels    1;  Gaps    1;

Qy         48 AGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAG 107
              ||| |  ||  |||  | |  |      | || |  ||||||||||||||||||| |||||
Db        436 AGAAACAGTGTGGCAGGCCGGGTGCGGTGGTGCACACCTGTAATCCCAGCACTTTGGGAG 377
```

```
Qy        108 GCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGA 167
              ||||||||||||||||| |||||| ||| ||| |||||||| ||||||||||||||   |
Db        376 GCCAAGGTGGGCGGATCGCCTGAGATCAGGAGTTCGAGACCAGCCTGGCCAACATGGCAA 317

Qy        168 AACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTC 227
              |||||| ||| |||||||||||||||| ||||||| | ||||||  ||||||  ||
Db        316 AACCCCATCTCTACTAAAAATACAAAAATTAGCTGGATGTTGTGGCAGGTGCCTGTAATC 257

Qy        228 CCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCAC 287
              |||||||||||||| | ||| | || || | | | | |||
Db        256 CCAGCTACTCAGGAGGCTGAGGCAGGAGAATCTCTTGAACCTGGGAGGCGGAGGTTGCAC 197

Qy        288 AGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
                | | | || ||||| ||||:| |||||||||| || || |||||| | |||||
Db        196 TCCAGCCTGGGAGA-CAGAGGAAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAGAAAAA 138

Qy        348 CCATAAGACA 357
                | || | |
Db        137 GAAGAAAAAA 128


RESULT 31
US-10-995-561-13274/c
; Sequence 13274, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13274
;   LENGTH: 415117
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(415117)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13274

  Query Match             38.4%;  Score 153.4;  DB 10;  Length 415117;
  Best Local Similarity   68.5%;  Pred. No. 1.6e-35;
  Matches  224;  Conservative    2;  Mismatches   98;  Indels    3;  Gaps    1;

Qy         39 ATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATCCCAGCA 98
              |||| |  |||  | |   |     | || |      |    | |||||||||:||| ||
Db     228512 ATATAAGAAAGTAAATAAAACAGGGGCCAGGTGCAGTGGCTCACGCCTGTAATMCCAACA 228453

Qy         99 CTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCA 158
              ||| | |||||| ||||||| ||||||| |||||||| ||| |||||||||  ||||||||
Db     228452 CTTTGAGAGGCCGAGGTGGGTGGATCACTTGAGGTCAGGAGTTCGAGACCAGCCTGGCCA 228393

Qy        159 ACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGGCACACA 218
              |||||||||||||||||| |||||||||||||||| ||||  | | ||||||||||||
Db     228392 ACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCAGTCGTGGTGGCACACA 228333

Qy        219 CCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGCAGAGTGAGC 278
              |||||| |||||  ||||| |||| |  || ||||| |||||  |||| |||
Db     228332 CCTGTAATCCCACTTACTCGGGAGGCAAAGGTTGCAATGAGCCCAGATTGCACCACTGCA 228273

Qy        279 CGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAA 338
              |   | |   |  ||||||   | ||| |  | |:|   |  ||||| || || |  ||
Db     228272 CTCCAGCCTGGGCGACAGAG---GGAGACTCCGTCTCAAACCAAAAAAAAAAAAAAGTAA 228216

Qy        339 ACAAAAAAACCATAAGACATTGTCCAT 365
              | |||  ||  | || |  || | |||
Db     228215 AGAAAGTAAATAAAATAATTTATTCAT 228189


RESULT 32
US-10-301-480-598906
; Sequence 598906, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 598906
;   LENGTH: 984
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-598906

  Query Match             38.3%;  Score 153;  DB 12;  Length 984;
  Best Local Similarity   71.5%;  Pred. No. 1.9e-36;
  Matches  201;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||| ||||||||| |||||||||||||||||||| |||||||||||| |||
Db        225 ATGCCTGTAATGCCAGCACTTTGGGAGGCCAAGGTGGGCGGACCACCTGAGGTCAGGAGT 284

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| ||||| |||||||||||||| | |||| ||||||||||||||||||| ||||
Db        285 TCGAGACCAGCCTGGACAACATGGTGAAACGCTGTCTCTACTAAAAATACAAAAATTAGC 344

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
                |||||||||    |||||||| |||||||||||||||| | ||||  | ||
Db        345 CAGGCATGGTGGTGGGCACCTGTAATCCCAGCTACTCAGGAGGCTGAGACAGGAGAATTA 404

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              | ||      ||||||| || |||||    |      |||  || || |      |
Db        405 CGTGAGGTTGCAGTGAGCTGAGATCACGCCACTGCACTCCAGCCTAGGGACAGAGAATGG 464

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGT 361
              |     |   || |||||| ||||||       | ||| ||
Db        465 CTCTGTCTCAAAAAAAAAAAAAAAAATTAGCTGGGCATAGT 505


RESULT 33
US-10-301-480-1212315
; Sequence 1212315, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1212315
;   LENGTH: 984
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1212315
```

```
  Query Match             38.3%;  Score 153;  DB 12;  Length 984;
  Best Local Similarity   71.5%;  Pred. No. 1.9e-36;
  Matches  201;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||| |||||||| |||||||||||||||||||| |||||||||||| |||
Db         225 ATGCCTGTAATGCCAGCACTTTGGGAGGCCAAGGTGGGCGGACCACCTGAGGTCAGGAGT 284

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               |||||||| ||||| |||||||||||||| | |||| |||||||||||||||||| ||||
Db         285 TCGAGACCAGCCTGGACAACATGGTGAAACGCTGTCTCTACTAAAAATACAAAAATTAGC 344

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||||||||    |||||||| |||||||||||||||||| | ||||  | ||
Db         345 CAGGCATGGTGGTGGGCACCTGTAATCCCAGCTACTCAGGAGGCTGAGACAGGAGAATTA 404

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               | ||      ||||||| || |||||     |         |||  || || |     |
Db         405 CGTGAGGTTGCAGTGAGCTGAGATCACGCCACTGCACTCCAGCCTAGGGACAGAGAATGG 464

Qy         321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGT 361
               |   .  || |||||| ||||||       | ||| ||
Db         465 CTCTGTCTCAAAAAAAAAAAAAAAAAATTAGCTGGGCATAGT 505


RESULT 34
US-09-764-891-7647/c
; Sequence 7647, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7647
;   LENGTH: 4388
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-891-7647

  Query Match             38.1%;  Score 152.4;  DB 3;  Length 4388;
  Best Local Similarity   76.5%;  Pred. No. 5.3e-36;
  Matches  202;  Conservative    1;  Mismatches   52;  Indels    9;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||||| |||||||| |||||||| |||  |||||||||| |||||||| |||
Db        1275 AAGCCTGTAATTCCAGCACTTTGGGAGGCCGAGGCCGGCGGATCACTTGAGGTCAGGAGT 1216

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               |||||||| |||||||||||||||||||| |||||| |||||||||||||||| ||||
Db        1215 TCGAGACCAGCCTGGCCAACATGGTGAAACTCCGTCTCTACTAAAAATACAAAAATTAGC 1156

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               |  | |||||||  | | |||||| |||||||||||| |||| ||||| |||||| ||||
Db        1155 CGCACGTGGTGGCGCGCGCCTGTAATCCCAGCTACTCGGGAGGCGGAGGTTGCAGCGAGC 1096

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               |||||||      |    |    |    |          || || |||||:||||||
Db        1095 AGAGATCGTGCCATTGCACTCCAGTCTGGGCGA---------CACAGCGAGACTCCGTCT 1045

Qy         321 CAAAAACAACAACAAAAAACAAAA 344
               |||||| || || |||||| ||||
Db        1044 CAAAAAAAAAAAAAAAAAAAAAAA 1021


RESULT 35
US-09-967-768A-314/c
; Sequence 314, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
;  APPLICANT: Augustus, Meena
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-72
;  CURRENT APPLICATION NUMBER: US/09/967,768A
;  CURRENT FILING DATE:  2001-09-28
;  PRIOR APPLICATION NUMBER: US/60/236,109
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US/60/236,034
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US/60/236,111
;  PRIOR FILING DATE: 2000-09-28
;  NUMBER OF SEQ ID NOS: 325
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 314
;   LENGTH: 174424
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-967-768A-314

  Query Match             38.0%;  Score 152;  DB 3;  Length 174424;
  Best Local Similarity   71.8%;  Pred. No. 3e-35;
  Matches  214;  Conservative    0;  Mismatches   80;  Indels    4;  Gaps    1;

Qy           6 TACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGG 65
               | |||||   | || ||  || |||    |      ||   |||||    | |||||  ||
Db       48438 TCCTCAGGTGTCTGGTGATCCTTGGTTGTCTTCTCATATTTAAGAATGGGGCAGGCTGGG 48379

Qy          66 CATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCA 125
               ||||  ||         |||||||||||||||| ||||| |||||||| | ||||||||||||
Db       48378 CATGGCA----GCTCATGCCTGTAATCCCAACACTTTGGGAGGCCGAAGTGGGCGGATCA 48323

Qy         126 CCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAA 185
               |||||||||||||||| || |||||||||||||||||||||||||| || | |||||||
Db       48322 CCTGAGGTCAAGAGATCAAGTCCATCCTGGCCAACATGGTGAAACCCTGTTTGTACTAAA 48263

Qy         186 AATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCG 245
               |||||||||| |||||||||||||||| || |||||||||||||||||||   ||||||
Db       48262 AATACAAAAATTAGCTGGGCATGGTGGTACGCACCTGTAGTCCCAGCTACTTGGGAGCCT 48203

Qy         246 GAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
               |||   | ||         ||   |.|   | |  ||   ||  || |  ||| || ||
Db       48202 GAGGCAGGAGAATCGCTTGAATCCGGGAGGTGGAGATTGCAGTGAGCCGAGATTGTGC 48145


RESULT 36
US-09-960-706-969/c
; Sequence 969, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplas
;  TITLE OF INVENTION:  Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-01US
;  CURRENT APPLICATION NUMBER: US/09/960,706
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: 60/223,323
;  PRIOR FILING DATE: 2000-08-07
;  PRIOR APPLICATION NUMBER: 09/873,319
;  PRIOR FILING DATE: 2001-06-05
;  NUMBER OF SEQ ID NOS: 1124
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 969
;   LENGTH: 174424
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134280A1 U52112
US-09-960-706-969

  Query Match             38.0%;  Score 152;  DB 3;  Length 174424;
  Best Local Similarity   71.8%;  Pred. No. 3e-35;
```

```
  Matches  214;  Conservative    0;  Mismatches   80;  Indels    4;  Gaps    1;

Qy          6 TACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGG 65
              | |||||    | || ||   || |||     |      ||   |||||     | |||||  ||
Db      48438 TCCTCAGGTGTCTGGTGATCCTTGGTTGTCTTCTCATATTTAAGAATGGGGCAGGCTGGG 48379

Qy         66 CATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCA 125
              ||||  ||         ||||||||||||||| |||||  |||||||| | ||||||||||||
Db      48378 CATGGCA----GCTCATGCCTGTAATCCCAACACTTTGGGAGGCCGAAGTGGGCGGATCA 48323

Qy        126 CCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAA 185
              |||||||||||||||| || |||||||||||||||||||||||| || | |||||||
Db      48322 CCTGAGGTCAAGAGATCAAGTCCATCCTGGCCAACATGGTGAAACCCTGTTTGTACTAAA 48263

Qy        186 AATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCG 245
              ||||||||||  |||||||||||||||| || ||||||||||||||||||||  ||||||
Db      48262 AATACAAAAATTAGCTGGGCATGGTGGTACGCACCTGTAGTCCCAGCTACTTGGGAGCCT 48203

Qy        246 GAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
              |||   | ||         ||   | |   | |  ||   ||  || |  ||| || ||
Db      48202 GAGGCAGGAGAATCGCTTGAATCCGGGAGGTGGAGATTGCAGTGAGCCGAGATTGTGC 48145


RESULT 37
US-10-843-641A-6459/c
; Sequence 6459, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
;  APPLICANT: Avalon Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-189
;  CURRENT APPLICATION NUMBER: US/10/843,641A
;  CURRENT FILING DATE:  2004-05-12
;  PRIOR APPLICATION NUMBER: US/09/873,367
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6459
;   LENGTH: 174424
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-843-641A-6459

  Query Match             38.0%;  Score 152;  DB 10;  Length 174424;
  Best Local Similarity   71.8%;  Pred. No. 3e-35;
  Matches  214;  Conservative    0;  Mismatches   80;  Indels    4;  Gaps    1;

Qy          6 TACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGG 65
              | |||||    | || ||   || |||     |      ||   |||||     | |||||  ||
Db      48438 TCCTCAGGTGTCTGGTGATCCTTGGTTGTCTTCTCATATTTAAGAATGGGGCAGGCTGGG 48379

Qy         66 CATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCA 125
              ||||  ||         ||||||||||||||| |||||  |||||||| | ||||||||||||
Db      48378 CATGGCA----GCTCATGCCTGTAATCCCAACACTTTGGGAGGCCGAAGTGGGCGGATCA 48323

Qy        126 CCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAA 185
              |||||||||||||||| || |||||||||||||||||||||||| || | |||||||
Db      48322 CCTGAGGTCAAGAGATCAAGTCCATCCTGGCCAACATGGTGAAACCCTGTTTGTACTAAA 48263

Qy        186 AATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCG 245
              ||||||||||  |||||||||||||||| || ||||||||||||||||||||  ||||||
Db      48262 AATACAAAAATTAGCTGGGCATGGTGGTACGCACCTGTAGTCCCAGCTACTTGGGAGCCT 48203

Qy        246 GAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
              |||   | ||         ||   | |   | |  ||   ||  || |  ||| || ||
Db      48202 GAGGCAGGAGAATCGCTTGAATCCGGGAGGTGGAGATTGCAGTGAGCCGAGATTGTGC 48145


RESULT 38
US-10-723-860-2392/c
; Sequence 2392, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Ginsburg, Wendy M.
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
;  TITLE OF INVENTION:  Methods for Screening for Soft Tissue Sarcoma Modulators
;  FILE REFERENCE: 05882.0193.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/723,860
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,739
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 8393
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2392
;   LENGTH: 181343
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-723-860-2392

  Query Match             38.0%;  Score 152;  DB 9;  Length 181343;
  Best Local Similarity   71.8%;  Pred. No. 3.1e-35;
  Matches  214;  Conservative    0;  Mismatches   80;  Indels    4;  Gaps    1;

Qy          6 TACTCAGCCATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGG 65
              | |||||    | || ||   || |||     |      ||   |||||     | |||||  ||
Db      48439 TCCTCAGGTGTCTGGTGATCCTTGGTTGTCTTCTCATATTTAAGAATGGGGCAGGCTGGG 48380

Qy         66 CATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCA 125
              ||||  ||         ||||||||||||||| |||||  |||||||| | ||||||||||||
Db      48379 CATGGCA----GCTCATGCCTGTAATCCCAACACTTTGGGAGGCCGAAGTGGGCGGATCA 48324

Qy        126 CCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAA 185
              |||||||||||||||| || |||||||||||||||||||||||| || | |||||||
Db      48323 CCTGAGGTCAAGAGATCAAGTCCATCCTGGCCAACATGGTGAAACCCTGTTTGTACTAAA 48264

Qy        186 AATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCG 245
              ||||||||||  |||||||||||||||| || ||||||||||||||||||||  ||||||
Db      48263 AATACAAAAATTAGCTGGGCATGGTGGTACGCACCTGTAGTCCCAGCTACTTGGGAGCCT 48204

Qy        246 GAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
              |||   | ||         ||   | |   | |  ||   ||  || |  ||| || ||
Db      48203 GAGGCAGGAGAATCGCTTGAATCCGGGAGGTGGAGATTGCAGTGAGCCGAGATTGTGC 48146


RESULT 39
US-10-756-149-2215/c
; Sequence 2215, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
;  TITLE OF INVENTION:  METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
;  FILE REFERENCE: file
```

```
;  CURRENT APPLICATION NUMBER: US/10/756,149
;  CURRENT FILING DATE:  2004-01-12
;  NUMBER OF SEQ ID NOS: 5818
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2215
;   LENGTH: 181343
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-10-756-149-2215

  Query Match             38.0%;  Score 152;  DB 10;  Length 181343;
  Best Local Similarity   71.8%;  Pred. No. 3.1e-35;
  Matches  214;  Conservative    0;  Mismatches   80;  Indels    4;  Gaps    1;

Qy          6 TACTCAGCCATGTGCTGGGCCATGGAACCCAAATATTAATAAGACATTGTCAGGCCAGG 65
              | |||||   | || ||  || |||      |      ||     |||||     | |||||  ||
Db      48439 TCCTCAGGTGTCTGGTGATCCTTGGTTGTCTTCTCATATTTAAGAATGGGGCAGGCTGGG 48380

Qy         66 CATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCA 125
              |||| ||    ||||||||||||| ||||| |||||||| | |||||||||||
Db      48379 CATGGCA----GCTCATGCCTGTAATCCCAACACTTTGGGAGGCCGAAGTGGGCGGATCA 48324

Qy        126 CCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAA 185
              |||||||||||||||| || ||||||||||||||||||||||||| || | |||||||
Db      48323 CCTGAGGTCAAGAGATCAAGTCCATCCTGGCCAACATGGTGAAACCCTGTTTGTACTAAA 48264

Qy        186 AATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCG 245
              ||||||||| ||||||||||||||| || ||||||||||||||||||  ||||||
Db      48263 AATACAAAAATTAGCTGGGCATGGTGGTACGCACCTGTAGTCCCAGCTACTTGGGAGCCT 48204

Qy        246 GAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
              |||  |||        ||    | |    | |  ||     ||  || |   ||| || ||
Db      48203 GAGGCAGGAGAATCGCTTGAATCCGGGAGGTGGAGATTGCAGTGAGCCGAGATTGTGC 48146


RESULT 40
US-09-925-065A-45529
; Sequence 45529, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45529
;   LENGTH: 599
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-45529

  Query Match             37.9%;  Score 151.6;  DB 4;  Length 599;
  Best Local Similarity   74.2%;  Pred. No. 4.2e-36;
  Matches  224;  Conservative    2;  Mismatches   56;  Indels   20;  Gaps    2;

Qy         83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
              |||||||||||||||||||| ||||||| |||||||||||||||||||||||||| ||| ||
Db          2 GCCTGTAATCCCAGCACTTTGGGAGGCTAAGGTGGGCGGATCACCTGAGGTCAGGAGTTC 61

Qy        143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
              |||||| |||||||||||||||||||||| |||| ||||||||||||||||||| |||| |
Db         62 GAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCCG 121

Qy        203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCC----GGAGATTGCAGTGA 258
              | ||||||||||| |||||||| |||||||||||  ||||  |      | |||  |     |||
Db        122 GACATGGTGGCACGCACCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGAATCACTTGA 181

Qy        259 GCTGAGATCGCAGAGTGAGCCGAAATCACAGATCA----------------CAGAGTGAG 302
               |   :|    ||||||    || |   | |   ||||||                |    |
Db        182 ACCCRGGAGGCAGAGGTTGCAGTGAGCCAAGATCACGCCACTGCACTTCAGCCTGGGTGA 241

Qy        303 CAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTC 362
              ||||||||||:|||||||||||||| ||   |||||   ||||||   | || |||    |
Db        242 CAGAGTGAGACTCCGTCTCAAAAACAAAAAACAAAAAACAAAAAAAAACAAAACAAAAAC 301

Qy        363 CA 364
              ||
Db        302 CA 303


RESULT 41
US-09-925-065A-45529
; Sequence 45529, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45529
;   LENGTH: 599
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-45529

  Query Match             37.9%;  Score 151.6;  DB 5;  Length 599;
  Best Local Similarity   74.2%;  Pred. No. 4.2e-36;
  Matches  224;  Conservative    2;  Mismatches   56;  Indels   20;  Gaps    2;

Qy         83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
              |||||||||||||||||||| ||||||| |||||||||||||||||||||||||| ||| ||
Db          2 GCCTGTAATCCCAGCACTTTGGGAGGCTAAGGTGGGCGGATCACCTGAGGTCAGGAGTTC 61

Qy        143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
              |||||| |||||||||||||||||||||| |||| ||||||||||||||||||| |||| |
Db         62 GAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCCG 121

Qy        203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCC----GGAGATTGCAGTGA 258
              | ||||||||||| |||||||| |||||||||||  ||||  |      | |||  |     |||
Db        122 GACATGGTGGCACGCACCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGAATCACTTGA 181

Qy        259 GCTGAGATCGCAGAGTGAGCCGAAATCACAGATCA----------------CAGAGTGAG 302
               |   :|    ||||||    || |   | |   ||||||                |    |
Db        182 ACCCRGGAGGCAGAGGTTGCAGTGAGCCAAGATCACGCCACTGCACTTCAGCCTGGGTGA 241

Qy        303 CAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTC 362
              ||||||||||:|||||||||||||| ||   |||||   ||||||   | || |||    |
Db        242 CAGAGTGAGACTCCGTCTCAAAAACAAAAAACAAAAAACAAAAAAAAACAAAACAAAAAC 301
```

```
Qy        363 CA 364
              ||
Db        302 CA 303


RESULT 42
US-10-301-480-146767
; Sequence 146767, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 146767
;   LENGTH: 599
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-146767

  Query Match             37.9%;  Score 151.6;  DB 12;  Length 599;
  Best Local Similarity   74.2%;  Pred. No. 4.2e-36;
  Matches  224;  Conservative    2;  Mismatches   56;  Indels   20;  Gaps    2;

Qy         83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
              |||||||||||||||||||| ||||||| |||||||||||||||||||||||||| ||| ||
Db          2 GCCTGTAATCCCAGCACTTTGGGAGGCTAAGGTGGGCGGATCACCTGAGGTCAGGAGTTC 61

Qy        143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
              ||||||| |||||||||||||||||||| |||| |||||||||||||||| |||| |
Db         62 GAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCCG 121

Qy        203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCC----GGAGATTGCAGTGA 258
              | |||||||||||| |||||||| |||||||||||  |||| |    | ||| |    |||
Db        122 GACATGGTGGCACGCACCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGAATCACTTGA 181

Qy        259 GCTGAGATCGCAGAGTGAGCCGAAATCACAGATCA----------------CAGAGTGAG 302
               |  :|   ||||||   || |  | |  ||||||                |   |
Db        182 ACCCRGGAGGCAGAGGTTGCAGTGAGCCAAGATCACGCCACTGCACTTCAGCCTGGGTGA 241

Qy        303 CAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTC 362
              |||||||||||:|||||||||||||| ||   |||||  ||||||  | || |||    |
Db        242 CAGAGTGAGACTCCGTCTCAAAAACAAAAAACAAAAAACAAAAAAAAACAAAACAAAAAC 301

Qy        363 CA 364
              ||
Db        302 CA 303


RESULT 43
US-10-301-480-760176
; Sequence 760176, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 760176
;   LENGTH: 599
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-760176

  Query Match             37.9%;  Score 151.6;  DB 12;  Length 599;
  Best Local Similarity   74.2%;  Pred. No. 4.2e-36;
  Matches  224;  Conservative    2;  Mismatches   56;  Indels   20;  Gaps    2;

Qy         83 GCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATC 142
              |||||||||||||||||||| ||||||| |||||||||||||||||||||||||| ||| ||
Db          2 GCCTGTAATCCCAGCACTTTGGGAGGCTAAGGTGGGCGGATCACCTGAGGTCAGGAGTTC 61

Qy        143 GAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTG 202
              ||||||| |||||||||||||||||||| |||| |||||||||||||||| |||| |
Db         62 GAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCCG 121

Qy        203 GGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCC----GGAGATTGCAGTGA 258
              | |||||||||||| |||||||| |||||||||||  |||| |    | ||| |    |||
Db        122 GACATGGTGGCACGCACCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGAATCACTTGA 181

Qy        259 GCTGAGATCGCAGAGTGAGCCGAAATCACAGATCA----------------CAGAGTGAG 302
               |  :|   ||||||   || |  | |  ||||||                |   |
Db        182 ACCCRGGAGGCAGAGGTTGCAGTGAGCCAAGATCACGCCACTGCACTTCAGCCTGGGTGA 241

Qy        303 CAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTC 362
              |||||||||||:|||||||||||||| ||   |||||  ||||||  | || |||    |
Db        242 CAGAGTGAGACTCCGTCTCAAAAACAAAAAACAAAAAACAAAAAAAAACAAAACAAAAAC 301

Qy        363 CA 364
              ||
Db        302 CA 303


RESULT 44
US-10-450-826-93/c
; Sequence 93, Application US/10450826
; Publication No. US20040101818A1
; GENERAL INFORMATION:
;  APPLICANT: Ji, Darren
;  APPLICANT:  Axelrod, Douglas W.
;  APPLICANT:  Cook, Jonathon S.
;  APPLICANT:  Jaiswal, Neelam
;  APPLICANT:  Eistein, Richard
;  APPLICANT:  Houghton, Adam
;  APPLICANT:  Mertz, Lawrence
;  TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiati
;  FILE REFERENCE: 044921-5039-WO
;  CURRENT APPLICATION NUMBER: US/10/450,826
;  CURRENT FILING DATE:  2003-06-18
;  PRIOR APPLICATION NUMBER: US 60/255,882
;  PRIOR FILING DATE: 2000-12-18
;  PRIOR APPLICATION NUMBER: US 60/285,691
;  PRIOR FILING DATE: 2001-04-24
;  NUMBER OF SEQ ID NOS: 149
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 93
;   LENGTH: 169739
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. AC005082
US-10-450-826-93

  Query Match             37.9%;  Score 151.4;  DB 8;  Length 169739;
  Best Local Similarity   70.9%;  Pred. No. 4.6e-35;
  Matches  229;  Conservative    1;  Mismatches   87;  Indels    6;  Gaps    2;

Qy         30 GGAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTA 89
              || |   |||| ||| | | | ||    ||    | || |       |   || |||||||
Db     125892 GGTAAGAAAATCTTAGTTAAAGATCTAAAGTTTGGCCAGGCACAGTGGCTAACGCCTGTA 125833
```

```
Qy          90 ATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCA 149
               ||||||||||||||||||||||||  || ||||||||||||||| ||| |||||||||
Db      125832 ATCCCAGCACTTCGGGAGGCCAAGGCAGGTGGATCACCTGAGGTCAGGAGTTCGAGACCA 125773

Qy         150 TCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGG 209
               |||||||||||||||||||||| |||| |||||||||||||||||| |||| ||||||||
Db      125772 GCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCCGGGCATGG 125713

Qy         210 TGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAGATCGC 269
               ||||| |||||||||||||||| ||||||| ||| | |||    | ||         || | |
Db      125712 TGGCAGACACCTGTAGTCCCAACTACTCAAGAGGCTGAGGCAGGAGAATCACTTGAAC-C 125654

Qy         270 AGAGTGAGCCGAAATCACA-----GATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAA 324
               | |||||  | | ||| ||     |  | |       |   |  |||||:|| |||| ||
Db      125653 TGGGTGAGTCAAGATCGCAACATGGCACTCCAGCCTGGGCAACAGAGACTCCATCTCCAA 125594

Qy         325 AACAACAACAAAAAACAAAAAAA 347
               || || ||  ||| | |||| ||
Db      125593 AAAAATAAATAAATAAAAAATAA 125571


RESULT 45
US-10-331-053-70
; Sequence 70, Application US/10331053
; Publication No. US20040197778A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc S. Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001100
;  CURRENT APPLICATION NUMBER: US/10/331,053
;  CURRENT FILING DATE:  2002-12-26
;  NUMBER OF SEQ ID NOS: 86
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 70
;   LENGTH: 133300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(133300)
;   OTHER INFORMATION: n = A,T,C or G
US-10-331-053-70

  Query Match             37.8%;  Score 151;  DB 9;  Length 133300;
  Best Local Similarity   77.3%;  Pred. No. 5.6e-35;
  Matches  211;  Conservative    1;  Mismatches   51;  Indels   10;  Gaps    2;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||||||||||||||| |||||||| ||| ||||||||||||||||||||| |||
Db        8926 ACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATCACCTGAGGTCAGGAGT 8985

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAA-TAG 199
               ||||||||| |||||||| |||||||||||||||||| |||||||||||||||||| |||
Db        8986 TCGAGACCAGCCTGGCCAAGATGGTGAAACCCCGTCTCTACTAAAAATACAAAAAATTAG 9045

Qy         200 CTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAG 259
               | |||| ||||||||  | |||||| |||||||||||| |||| | ||| ||||||||||
Db        9046 CCGGGCGTGGTGGCAGGCGCCTGTAATCCCAGCTACTCGGGAGGCAGAGGTTGCAGTGAG 9105

Qy         260 CTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTC 319
               | ||||||||      | | |   |   |         ||||||||  |:|||||
Db        9106 CCGAGATCGCGTCACTGCACTCCAGCCTGGGCGA---------CAGAGTGAAGCTCCGTC 9156

Qy         320 TCAAAAACAACAACAAAAAACAAAAAAACCATA 352
               ||| | | || ||  ||||| ||||||   |||
Db        9157 TCAGAGAAAAAAAAGAAAAAAAAAAAAGAAATA 9189


RESULT 46
US-10-995-561-13489/c
; Sequence 13489, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13489
;   LENGTH: 195998
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(195998)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13489

  Query Match             37.8%;  Score 151;  DB 10;  Length 195998;
  Best Local Similarity   76.8%;  Pred. No. 6.5e-35;
  Matches  199;  Conservative    1;  Mismatches   51;  Indels    8;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||||||||||||||| ||||  || ||| ||| || |||||||||||||| |||
Db      141321 ACGCCTGTAATCCCAGCACTTTGGGAAACCGAGGAGGGTGGGTCACCTGAGGTCAGGAGT 141262

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               |||| |||| |||||||||||||||||||||||||||| ||  |||| |||||||| ||||
Db      141261 TCGACACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTATGAAAAGTACAAAAACTAGC 141202

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||||||||||   | |||||| |||||||||||| ||| | ||| ||||||||| |
Db      141201 CGGGCATGGTGGCGGGCGCCTGTAATCCCAGCTACTCACGAGGCAGAGGTTGCAGTGAAC 141142

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               || ||||          | |   |||       | | ||||   ||| |||||:||||||
Db      141141 CGAAATCG--------TGCCACTGCACTCCAGCCTGGGTGACGTGAGCGAGACTCCGTCT 141090

Qy         321 CAAAAACAACAACAAAAAA 339
               |||||| || || ||||||
Db      141089 CAAAAAAAAAAAAAAAAAA 141071


RESULT 47
US-09-764-860-797/c
; Sequence 797, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC008
;  CURRENT APPLICATION NUMBER: US/09/764,860
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1198
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 797
;   LENGTH: 32146
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-860-797

  Query Match             37.7%;  Score 150.6;  DB 3;  Length 32146;
  Best Local Similarity   68.5%;  Pred. No. 4.2e-35;
  Matches  224;  Conservative    1;  Mismatches   95;  Indels    7;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | ||||||||| | ||||||| |||||||||||| ||| ||||||| |||||||| |||
Db       2437 ACGCCTGTAATGCTAGCACTTTGGGAGGCCAAGGCGGGAGGATCACTTGAGGTCAGGAGT 2378
```

```
Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||| | ||||||||||||||||||||||| |||| ||||||||||||||||| ||||
Db       2377 TTGAGACAAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGC 2318

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAG-------CCGGAGATTGC 253
              |||||||||||||  ||||||||||||||||| ||  ||||       | ||||| |
Db       2317 AGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTGCTTGGGAGACTGAGCCAGGAGAATTG 2258

Qy        254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGAC 313
              ||| |   |   || |||   ||    | |  ||||   |        |    |||||
Db       2257 CTTGAACCCGGGAGGCGGAGGTTGCAATGAGCCGAGATTGTACTACTGCAATGCAGAGAC 2198

Qy        314 KCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTT 373
              :||||||||||||| || || |||||| |||||||  |||| |        ||   | ||
Db       2197 TCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGATAAAAGTAGAGACAATAGGGGGA 2138

Qy        374 CCCAGACTATTGCAGGAGACCAAAAAG 400
              |    | | | | |||    | ||||
Db       2137 TCTCAGCAAATACTGGATTTAACAAAG 2111
```

RESULT 48
US-10-074-095-797/c
; Sequence 797, Application US/10074095
; Publication No. US20030077704A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C1
; CURRENT APPLICATION NUMBER: US/10/074,095
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: 09/764,860
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681

; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match 37.7%; Score 150.6; DB 6; Length 32146;
Best Local Similarity 68.5%; Pred. No. 4.2e-35;
Matches 224; Conservative 1; Mismatches 95; Indels 7; Gaps 1;

```
Qy        81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
             | ||||||||| | ||||||| |||||||||||| ||| ||||||| |||||||| |||
Db      2437 ACGCCTGTAATGCTAGCACTTTGGGAGGCCAAGGCGGGAGGATCACTTGAGGTCAGGAGT 2378

Qy       141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
             | ||||| | |||||||||||||||||||||| |||| |||||||||||||||| ||||
Db      2377 TTGAGACAAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGC 2318

Qy       201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAG-------CCGGAGATTGC 253
             |||||||||||||   |||||||||||||||||| ||  ||||       | ||||| |
Db      2317 AGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTGCTTGGGAGACTGAGCCAGGAGAATTG 2258

Qy       254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGAC 313
             ||| |   |   || |||  ||   | |  |||| |       |      |||||
Db      2257 CTTGAACCCGGGAGGCGGAGGTTGCAATGAGCCGAGATTGTACTACTGCAATGCAGAGAC 2198

Qy       314 KCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTT 373
             :|||||||||||| || || |||||| |||||||  |||| |       ||   | ||
Db      2197 TCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAGATAAAAGTAGAGACAATAGGGGGA 2138

Qy       374 CCCAGACTATTGCAGGAGACCAAAAAG 400
             |   | | | | |||   | ||||
Db      2137 TCTCAGCAAATACTGGATTTAACAAAG 2111
```

RESULT 49
US-10-212-872-797/c
; Sequence 797, Application US/10212872
; Publication No. US20030215893A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C2
; CURRENT APPLICATION NUMBER: US/10/212,872
; CURRENT FILING DATE: 2002-08-07
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 797
; LENGTH: 32146
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-212-872-797

Query Match 37.7%; Score 150.6; DB 7; Length 32146;
Best Local Similarity 68.5%; Pred. No. 4.2e-35;
Matches 224; Conservative 1; Mismatches 95; Indels 7; Gaps 1;

```
Qy        81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
             | ||||||||| | ||||||| |||||||||||| ||| ||||||| |||||||| |||
Db      2437 ACGCCTGTAATGCTAGCACTTTGGGAGGCCAAGGCGGGAGGATCACTTGAGGTCAGGAGT 2378

Qy       141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
             | ||||| | |||||||||||||||||||||| |||| |||||||||||||||| ||||
Db      2377 TTGAGACAAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGC 2318

Qy       201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAG-------CCGGAGATTGC 253
             |||||||||||||   |||||||||||||||||| ||  ||||       | ||||| |
Db      2317 AGGGCATGGTGGCGGGCACCTGTAGTCCCAGCTGCTTGGGAGACTGAGCCAGGAGAATTG 2258

Qy       254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGAC 313
             ||| |   |   || |||  ||   | |  |||| |       |      |||||
Db      2257 CTTGAACCCGGGAGGCGGAGGTTGCAATGAGCCGAGATTGTACTACTGCAATGCAGAGAC 2198
```

```
Qy        314 KCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTT 373
              :|||||||||||| || || |||||| |||||||  |||| |        ||   | ||
Db       2197 TCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAGATAAAAGTAGAGACAATAGGGGGA 2138

Qy        374 CCCAGACTATTGCAGGAGACCAAAAAG 400
               |  .  | | | | |||    | ||||
Db       2137 TCTCAGCAAATACTGGATTTAACAAAG 2111


RESULT 50
US-09-893-348-9/c
; Sequence 9, Application US/09893348
; Patent No. US20020072493A1
; GENERAL INFORMATION:
;  APPLICANT: EISENBACH-SCHWARTZ, Michal
;  APPLICANT:  COHEN, Irun R.
;  APPLICANT:  BESERMAN, Pierre
;  APPLICANT:  MOSONEGO, Alon
;  APPLICANT:  MOALEM, Gila
;  TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND THEIR USE
;  FILE REFERENCE: EIS-SCHWARTZ=2A
;  CURRENT APPLICATION NUMBER: US/09/893,348
;  CURRENT FILING DATE:  2001-06-28
;  PRIOR APPLICATION NUMBER: US 09/314,161
;  PRIOR FILING DATE: 1999-05-19
;  PRIOR APPLICATION NUMBER: US 09/218,277
;  PRIOR FILING DATE: 1998-12-22
;  PRIOR APPLICATION NUMBER: PCT/US98/14715
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: IL 124500
;  PRIOR FILING DATE: 1998-05-19
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
;  LENGTH: 17538
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-893-348-9

  Query Match             37.6%;  Score 150.4;  DB 3;  Length 17538;
  Best Local Similarity   72.1%;  Pred. No. 3.8e-35;
  Matches  196;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| || |||||||||  ||| ||||||||||||||| | ||
Db      13528 ATGCCTGTAATCCCAGCACTTTGGAAGGCCAAGGCAGGCAGATCACCTGAGGTCAGGGGT 13469

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || |||||| |||||||||||||||||||||||||| |||||||||||||||| ||||
Db      13468 TCAAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 13409

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||||||||||||   | |||||||| ||||| |||  |||| | |||    | ||
Db      13408 TGGGCATGGTGGCATGCGCCTGTAGTTCCAGCAACTTGGGAGGCTGAGGCAGGAGAATCG 13349

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                || | |||   | |  |    ||  || |  ||| || ||      |  |    ||
Db      13348 CTTGAACCCAGGAGGTGGAGGTTGCAGTGAGCCGAGATTGTGCCATTGCACTCCATCTCA 13289

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATA 352
               ||||||  || |  |||| |||| ||  | |
Db      13288 AAAAAACTTCATCTCAAAAAAAAAGAAAGAAA 13257


Search completed: July 17, 2006, 22:44:38
Job time : 1466 secs
```

GenCore version 5.1.9

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 21:21:05 ; Search time 407 Seconds
(without alignments)
1370.274 Million cell updates/sec

Title: SEQ1-47502C
Perfect score: 399.6
Sequence: 1 ccaggtactcagccatgtgc.....tattgcaggagaccaaaaag 400

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 886355 seqs, 697127050 residues

Total number of hits satisfying chosen parameters: 1772710

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : Published_Applications_NA_New:*
1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq1:*
9: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	155.8	39.0	216387	8	US-11-266-748A-23004	Sequence 23004, A
c	2	151	37.8	1000	8	US-11-266-748A-204137	Sequence 204137,
	3	149.4	37.4	1000	8	US-11-266-748A-391815	Sequence 391815,
c	4	149.4	37.4	1000	8	US-11-266-748A-482533	Sequence 482533,
	5	147.8	37.0	1000	8	US-11-266-748A-118317	Sequence 118317,
c	6	147.8	37.0	1000	8	US-11-266-748A-160481	Sequence 160481,
	7	147.8	37.0	1000	8	US-11-266-748A-403127	Sequence 403127,
c	8	147.8	37.0	1000	8	US-11-266-748A-474173	Sequence 474173,
	9	147.6	36.9	11820	8	US-11-319-952-56	Sequence 56, Appl
	10	147.2	36.8	897	8	US-11-266-748A-219455	Sequence 219455,
c	11	147.2	36.8	897	8	US-11-266-748A-239441	Sequence 239441,
c	12	147	36.8	2909	8	US-11-293-697-1113	Sequence 1113, Ap
c	13	146.2	36.6	112595	8	US-11-266-748A-23290	Sequence 23290, A
c	14	145.6	36.4	6372	8	US-11-266-748A-30503	Sequence 30503, A
c	15	145	36.3	3324	8	US-11-266-748A-24156	Sequence 24156, A
	16	144.8	36.2	1000	8	US-11-266-748A-390100	Sequence 390100,
c	17	144.8	36.2	1000	8	US-11-266-748A-480818	Sequence 480818,
	18	144.4	36.1	227968	8	US-11-266-748A-60135	Sequence 60135, A
c	19	144.2	36.1	1000	8	US-11-266-748A-206136	Sequence 206136,
c	20	144.2	36.1	227968	8	US-11-266-748A-60135	Sequence 60135, A
c	21	143	35.8	1000	8	US-11-266-748A-198112	Sequence 198112,
c	22	143	35.8	96217	8	US-11-266-748A-59101	Sequence 59101, A
	23	142.6	35.7	1945	8	US-11-266-748A-27021	Sequence 27021, A
	24	142.6	35.7	1945	8	US-11-266-748A-58369	Sequence 58369, A
	25	142.6	35.7	1421559	8	US-11-266-748A-28208	Sequence 28208, A
c	26	142.4	35.6	1000	8	US-11-266-748A-223929	Sequence 223929,
c	27	142.4	35.6	1000	8	US-11-266-748A-292231	Sequence 292231,
	28	142.4	35.6	1000	8	US-11-266-748A-293957	Sequence 293957,
	29	142.4	35.6	1000	8	US-11-266-748A-343660	Sequence 343660,
c	30	142.4	35.6	1000	8	US-11-266-748A-345386	Sequence 345386,
c	31	142.4	35.6	1000	8	US-11-266-748A-403868	Sequence 403868,
	32	142.4	35.6	1000	8	US-11-266-748A-406117	Sequence 406117,
	33	142.4	35.6	1000	8	US-11-266-748A-474914	Sequence 474914,
c	34	142.4	35.6	1000	8	US-11-266-748A-477163	Sequence 477163,
	35	142.2	35.6	157866	7	US-11-191-644-1	Sequence 1, Appli
c	36	142	35.5	2800	8	US-11-293-697-23	Sequence 23, Appl
c	37	142	35.5	3252	8	US-11-293-697-604	Sequence 604, App
	38	141.6	35.4	1000	8	US-11-266-748A-202571	Sequence 202571,
c	39	141.6	35.4	216387	8	US-11-266-748A-23004	Sequence 23004, A
	40	141.4	35.4	632	8	US-11-266-748A-201568	Sequence 201568,
c	41	141.4	35.4	2445	8	US-11-266-748A-59943	Sequence 59943, A
c	42	141.4	35.4	6521	6	US-10-517-441-32	Sequence 32, Appl
	43	141.4	35.4	86654	8	US-11-266-748A-60244	Sequence 60244, A
c	44	141.4	35.4	113853	8	US-11-266-748A-23291	Sequence 23291, A
c	45	141.4	35.4	154394	8	US-11-266-748A-58517	Sequence 58517, A
c	46	141.4	35.4	164429	8	US-11-266-748A-22662	Sequence 22662, A
c	47	141.2	35.3	86215	8	US-11-266-748A-24937	Sequence 24937, A
	48	141	35.3	1000	8	US-11-266-748A-119261	Sequence 119261,
c	49	141	35.3	1000	8	US-11-266-748A-161425	Sequence 161425,
	50	141	35.3	1000	8	US-11-266-748A-407602	Sequence 407602,
c	51	141	35.3	1000	8	US-11-266-748A-478648	Sequence 478648,
c	52	140.8	35.2	1000	8	US-11-266-748A-203363	Sequence 203363,
c	53	140.6	35.2	813	8	US-11-266-748A-214586	Sequence 214586,
	54	140.6	35.2	813	8	US-11-266-748A-237142	Sequence 237142,
c	55	140.6	35.2	1000	8	US-11-266-748A-203963	Sequence 203963,
	56	140.4	35.1	1000	8	US-11-266-748A-196966	Sequence 196966,
c	57	140.4	35.1	1000	8	US-11-266-748A-197423	Sequence 197423,
	58	140.4	35.1	2775	8	US-11-293-697-340	Sequence 340, App
	59	140.2	35.1	1000	8	US-11-266-748A-196084	Sequence 196084,
c	60	140	35.0	1000	8	US-11-266-748A-199445	Sequence 199445,
c	61	140	35.0	1000	8	US-11-266-748A-294379	Sequence 294379,
	62	140	35.0	1000	8	US-11-266-748A-345808	Sequence 345808,
c	63	140	35.0	1000	8	US-11-266-748A-406642	Sequence 406642,
	64	140	35.0	1000	8	US-11-266-748A-477688	Sequence 477688,
c	65	140	35.0	2012	8	US-11-266-748A-25833	Sequence 25833, A
c	66	140	35.0	6339	8	US-11-266-748A-56164	Sequence 56164, A
	67	140	35.0	59753	8	US-11-266-748A-25122	Sequence 25122, A
	68	139.8	35.0	1000	8	US-11-266-748A-197052	Sequence 197052,
c	69	139.8	35.0	1000	8	US-11-266-748A-204215	Sequence 204215,
c	70	139.8	35.0	1000	8	US-11-266-748A-205230	Sequence 205230,
	71	139.8	35.0	1000	8	US-11-266-748A-393904	Sequence 393904,
c	72	139.8	35.0	1000	8	US-11-266-748A-464950	Sequence 464950,
c	73	139.8	35.0	1769	8	US-11-266-748A-23096	Sequence 23096, A
c	74	139.8	35.0	1769	8	US-11-266-748A-60917	Sequence 60917, A
c	75	139.8	35.0	1985	8	US-11-266-748A-59034	Sequence 59034, A
	76	139.8	35.0	2507	8	US-11-293-697-2398	Sequence 2398, Ap
	77	139.8	35.0	40850	8	US-11-266-748A-22340	Sequence 22340, A
c	78	139.8	35.0	49525	8	US-11-266-748A-24903	Sequence 24903, A
c	79	139.8	35.0	84539	7	US-11-191-644-6	Sequence 6, Appli
c	80	139.8	35.0	84539	8	US-11-266-748A-31827	Sequence 31827, A
	81	139.8	35.0	124048	8	US-11-266-748A-59707	Sequence 59707, A
	82	139.8	35.0	126259	8	US-11-266-748A-61158	Sequence 61158, A
c	83	139.6	34.9	31857	8	US-11-266-748A-22572	Sequence 22572, A
c	84	139.6	34.9	171653	8	US-11-266-748A-22660	Sequence 22660, A
c	85	139.4	34.9	4728	8	US-11-266-748A-26985	Sequence 26985, A
	86	139.4	34.9	112595	8	US-11-266-748A-23290	Sequence 23290, A
c	87	139.2	34.8	1000	8	US-11-266-748A-205571	Sequence 205571,
	88	139.2	34.8	4918	8	US-11-266-748A-28411	Sequence 28411, A
	89	139.2	34.8	81914	8	US-11-266-748A-59116	Sequence 59116, A
	90	139.2	34.8	127145	8	US-11-266-748A-59222	Sequence 59222, A
c	91	139.2	34.8	242825	8	US-11-266-748A-32745	Sequence 32745, A
	92	139	34.8	5387	6	US-10-517-441-111	Sequence 111, App
	93	139	34.8	88029	8	US-11-270-796-20	Sequence 20, Appl
c	94	139	34.8	155379	8	US-11-266-748A-24170	Sequence 24170, A
c	95	139	34.8	155379	8	US-11-266-748A-59410	Sequence 59410, A
c	96	138.8	34.7	606	8	US-11-266-748A-211296	Sequence 211296,
	97	138.8	34.7	606	8	US-11-266-748A-235003	Sequence 235003,

c	98	138.8	34.7	13076	6	US-10-517-441-74 Sequence 74, Appl
	99	138.8	34.7	122568	8	US-11-266-748A-23292 Sequence 23292, A
c	100	138.8	34.7	124074	8	US-11-266-748A-28242 Sequence 28242, A
	101	138.8	34.7	183797	8	US-11-266-748A-22659 Sequence 22659, A
	102	138.6	34.7	754	8	US-11-266-748A-27577 Sequence 27577, A
	103	138.6	34.7	1000	8	US-11-266-748A-195150 Sequence 195150,
	104	138.6	34.7	1000	8	US-11-266-748A-207224 Sequence 207224,
	105	138.6	34.7	2460	8	US-11-293-697-595 Sequence 595, App
	106	138.6	34.7	9859	6	US-10-517-441-105 Sequence 105, App
	107	138.6	34.7	54550	8	US-11-318-813-42 Sequence 42, Appl
	108	138.6	34.7	113349	8	US-11-266-748A-22661 Sequence 22661, A
	109	138.6	34.7	189056	8	US-11-266-748A-29044 Sequence 29044, A
	110	138.6	34.7	340000	8	US-11-266-748A-61353 Sequence 61353, A
c	111	138.4	34.6	1000	8	US-11-266-748A-198430 Sequence 198430,
	112	138.4	34.6	1000	8	US-11-266-748A-200218 Sequence 200218,
c	113	138.2	34.6	1000	8	US-11-266-748A-201007 Sequence 201007,
	114	138.2	34.6	1000	8	US-11-266-748A-202270 Sequence 202270,
	115	138.2	34.6	1000	8	US-11-266-748A-204489 Sequence 204489,
	116	138.2	34.6	1000	8	US-11-266-748A-204653 Sequence 204653,
	117	138.2	34.6	1000	8	US-11-266-748A-281697 Sequence 281697,
c	118	138.2	34.6	1000	8	US-11-266-748A-308337 Sequence 308337,
	119	138.2	34.6	3010	8	US-11-293-697-1125 Sequence 1125, Ap
c	120	138.2	34.6	3315	8	US-11-266-748A-27139 Sequence 27139, A
c	121	138.2	34.6	13172	8	US-11-266-748A-56186 Sequence 56186, A
c	122	138.2	34.6	101500	8	US-11-266-748A-61740 Sequence 61740, A
	123	138.2	34.6	185035	8	US-11-266-748A-59932 Sequence 59932, A
	124	138.2	34.6	275434	8	US-11-266-748A-23986 Sequence 23986, A
c	125	138.2	34.6	275434	8	US-11-266-748A-23986 Sequence 23986, A
c	126	138.2	34.6	1421559	8	US-11-266-748A-28208 Sequence 28208, A
c	127	138	34.5	1000	8	US-11-266-748A-199736 Sequence 199736,
c	128	138	34.5	86081	8	US-11-266-748A-23657 Sequence 23657, A
c	129	138	34.5	88029	8	US-11-270-796-20 Sequence 20, Appl
c	130	138	34.5	106328	8	US-11-266-748A-25129 Sequence 25129, A
c	131	137.8	34.5	4471	8	US-11-266-748A-27527 Sequence 27527, A
c	132	137.8	34.5	4471	8	US-11-266-748A-27528 Sequence 27528, A
	133	137.8	34.5	68689	8	US-11-266-748A-23956 Sequence 23956, A
c	134	137.6	34.4	1000	8	US-11-266-748A-199869 Sequence 199869,
c	135	137.6	34.4	1000	8	US-11-266-748A-220450 Sequence 220450,
c	136	137.6	34.4	1000	8	US-11-266-748A-405873 Sequence 405873,
	137	137.6	34.4	1000	8	US-11-266-748A-476919 Sequence 476919,
c	138	137.6	34.4	2065	8	US-11-293-697-1789 Sequence 1789, Ap
c	139	137.6	34.4	5926	8	US-11-266-748A-22592 Sequence 22592, A
c	140	137.6	34.4	23722	8	US-11-266-748A-25108 Sequence 25108, A
	141	137.6	34.4	130263	8	US-11-266-748A-58551 Sequence 58551, A
	142	137.4	34.4	1000	8	US-11-266-748A-118098 Sequence 118098,
c	143	137.4	34.4	1000	8	US-11-266-748A-160262 Sequence 160262,
	144	137.4	34.4	1000	8	US-11-266-748A-220429 Sequence 220429,
	145	137.4	34.4	1000	8	US-11-266-748A-291058 Sequence 291058,
c	146	137.4	34.4	1000	8	US-11-266-748A-342487 Sequence 342487,
	147	137.4	34.4	1000	8	US-11-266-748A-402434 Sequence 402434,
c	148	137.4	34.4	1000	8	US-11-266-748A-473480 Sequence 473480,
	149	137.4	34.4	2630	8	US-11-266-748A-30772 Sequence 30772, A
	150	137.4	34.4	2630	8	US-11-266-748A-57101 Sequence 57101, A

ALIGNMENTS

```
RESULT 1
US-11-266-748A-23004
; Sequence 23004, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23004
;   LENGTH: 216387
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-23004

  Query Match             39.0%;  Score 155.8;  DB 8;  Length 216387;
  Best Local Similarity   77.0%;  Pred. No. 9.5e-24;
  Matches  218;  Conservative    1;  Mismatches   53;  Indels   11;  Gaps    2;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||| |||||||| |||||||| ||||||||||||||||||||
Db      195538 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACCTGAGGTCAAGAGA 195597

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               ||||||| |||||||||||||||||||||||||||||||||| |||||||||||||||| ||||
Db      195598 TCGAGACTATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 195657

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-------GCCGGAGATTGC 253
               ||||| |||||| |  || |||||||||||||||||||| |||       ||  |||| |
Db      195658 TGGGCGTGGTGGTATGCATCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAAGAGAGTTG 195717

Qy         254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGAT----CACAGAGTGAGCAGAGTG 309
                 ||| |        |||||| | || |  | |  ||||    ||          || || |
Db      195718 CTTGAACCTGAGAGGCAGAGGGTGCAGTGAACCAAGATACTCCAGCCTAGTGACACAGCG 195777

Qy         310 AGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATA 352
               ||||:|||||||||||| || || |||||| |||||| | ||
Db      195778 AGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAACACCTA 195820


RESULT 2
US-11-266-748A-204137/c
; Sequence 204137, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
```

```
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 204137
;    LENGTH: 1000
;    TYPE: DNA
;    ORGANISM: Homo Sapiens
US-11-266-748A-204137

  Query Match             37.8%;  Score 151;  DB 8;  Length 1000;
  Best Local Similarity   74.4%;  Pred. No. 7.2e-23;
  Matches  206;  Conservative    1;  Mismatches   61;  Indels    9;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||||| |||||||   |||    | | |||||||  | |||| |||
Db         303 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAAGTGCATCACCTAAAGTCACGAGT 244

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               || |||||| ||||||||||||||  |||| | |||| | ||||||||||||||| ||||
Db         243 TCAAGACCAGCCTGGCCAACATGGCAAAACTCTGTCTCTGCTAAAAATACAAAAATTAGC 184

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
                 ||||| ||||||  |||||||| |||||||||||| |||| ||||| ||||||||||
Db         183 CAGGCATAGTGGCAGGCACCTGTAATCCCAGCTACTCTGGAGGCGGAGGTTGCAGTGAGC 124

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               |||||| |||          |   | |   | | |          |||||  ||||:| ||||
Db         123 TGAGATTGCACCACTGCTCTTTAGCCTGGGTGA---------CAGAGCAAGACTCTGTCT 73

Qy         321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
               |||||| || || |||||| ||  |||  | || |||
Db          72 CAAAAAAAAAAAAAAAAAAAAAGCAAAAAACAAAACA 36


RESULT 3
US-11-266-748A-391815
; Sequence 391815, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 391815
;    LENGTH: 1000
;    TYPE: DNA
;    ORGANISM: Homo Sapiens
US-11-266-748A-391815

  Query Match             37.4%;  Score 149.4;  DB 8;  Length 1000;
  Best Local Similarity   74.0%;  Pred. No. 1.6e-22;
  Matches  205;  Conservative    1;  Mismatches   62;  Indels    9;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||| |||||| ||| |||||| |||||| |||| ||| |||
Db         732 ACGCCTGTAATCCCAGCACTTTGGGAGGTCAACGTGGGCAGATCACTTGAGTTCAGGAGT 791

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               || ||| || |||||||||||||||||||||||| | |||||||||||||||| ||||
Db         792 TCAAGAACAGCCTGGCCAACATGGTGAAACCCCGTATCTACTAAAAATACAAAAATTAGC 851

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
                 ||||||||||  || |||||  |||||||| ||||| | ||| ||||||||||
Db         852 CATGCATGGTGGCATGCATCTGTAACACCAGCTACTAAGGAGGCAGAGGTTGCAGTGAGC 911

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
               |||| ||       |   |   |          |   |   ||| ||  ||||:|| |||
Db         912 TGAGGTCATGCCATTGCACTTCA---------GCCTGGGCTGCACAGCCAGACTCCATCT 962

Qy         321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
               |||||| || || |||||| |||||||  | || | |
Db         963 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 999


RESULT 4
US-11-266-748A-482533/c
; Sequence 482533, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 482533
;    LENGTH: 1000
;    TYPE: DNA
;    ORGANISM: Homo Sapiens
US-11-266-748A-482533

  Query Match             37.4%;  Score 149.4;  DB 8;  Length 1000;
  Best Local Similarity   74.0%;  Pred. No. 1.6e-22;
  Matches  205;  Conservative    1;  Mismatches   62;  Indels    9;  Gaps    1;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||||||||||||||| |||||| ||| |||||| |||||| |||| ||| |||
Db         269 ACGCCTGTAATCCCAGCACTTTGGGAGGTCAACGTGGGCAGATCACTTGAGTTCAGGAGT 210

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               || ||| || |||||||||||||||||||||||| | |||||||||||||||| ||||
Db         209 TCAAGAACAGCCTGGCCAACATGGTGAAACCCCGTATCTACTAAAAATACAAAAATTAGC 150
```

```
Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||||||||| || |||||   ||||||||| ||||| | ||| ||||||||||||
Db        149 CATGCATGGTGGCATGCATCTGTAACACCAGCTACTAAGGAGGCAGAGGTTGCAGTGAGC 90

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              |||| ||       |   |  |          |   |   ||| ||  ||||:|| |||
Db         89 TGAGGTCATGCCATTGCACTTCA---------GCCTGGGCTGCACAGCCAGACTCCATCT 39

Qy        321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
              |||||| || || |||||| |||||||  | || | |
Db         38 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2


RESULT 5
US-11-266-748A-118317
; Sequence 118317, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 118317
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-118317

  Query Match           37.0%;  Score 147.8;  DB 8;  Length 1000;
  Best Local Similarity   68.4%;  Pred. No. 3.4e-22;
  Matches  221;  Conservative    1;  Mismatches   93;  Indels    8;  Gaps    1;

Qy         33 ACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATC 92
              || ||| | |  || |   | |     ||  | ||||       |    | ||||||||||
Db        677 ACACAATTTTACATTAAGAAATACTGTGCAGGCCATGCGTGGTGGCTCAGGCCTGTAATC 736

Qy         93 CCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCC 152
              ||||||||| |||||||| |||||||| |||||||| ||| |||||||||  ||
Db        737 CCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACCGGAGGTCAGGAGTTCGAGACCAGCC 796

Qy        153 TGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGG 212
              | |||||||| |||||||| |||||||||||||||||||||||| |||| ||||||||||
Db        797 TTGCCAACATAGTGAAACCCTGTCTTTACTAAAAATACAAAAATTAGCCGGGCATGGTGG 856

Qy        213 CACACACCTGTAGTCCCAGCTACTCAGGAG--------CCGGAGATTGCAGTGAGCTGAG 264
              ||  |||||||| |||||| ||||  ||||       |  ||   |||| |    ||
Db        857 CAGGCACCTGTAATCCCAGGTACTAGGGAGGCTTTTGAACCCAGGAGGCAGAGGTTGCAG 916

Qy        265 ATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAA 324
                 || |||   ||   | |          |   |     ||||||||| :| ||||||||
Db        917 CGAGCTGAGATCGCGCCACTGCACTCCAGCCTGGGTGATAGAGTGAGATTCAGTCTCAAA 976

Qy        325 AACAACAACAAAAAACAAAAAAA 347
              || || || |||||| |||||||
Db        977 AAAAAAAAAAAAAAAAAAAAAAA 999


RESULT 6
US-11-266-748A-160481/c
; Sequence 160481, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 160481
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-160481

  Query Match           37.0%;  Score 147.8;  DB 8;  Length 1000;
  Best Local Similarity   68.4%;  Pred. No. 3.4e-22;
  Matches  221;  Conservative    1;  Mismatches   93;  Indels    8;  Gaps    1;

Qy         33 ACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATC 92
              || ||| | |  || |   | |     ||  | ||||       |    | ||||||||||
Db        324 ACACAATTTTACATTAAGAAATACTGTGCAGGCCATGCGTGGTGGCTCAGGCCTGTAATC 265

Qy         93 CCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCC 152
              ||||||||| |||||||| |||||||| |||||||| ||| |||||||||  ||
Db        264 CCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACCGGAGGTCAGGAGTTCGAGACCAGCC 205

Qy        153 TGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGG 212
              | |||||||| |||||||| |||||||||||||||||||||||| |||| ||||||||||
Db        204 TTGCCAACATAGTGAAACCCTGTCTTTACTAAAAATACAAAAATTAGCCGGGCATGGTGG 145

Qy        213 CACACACCTGTAGTCCCAGCTACTCAGGAG--------CCGGAGATTGCAGTGAGCTGAG 264
              ||  |||||||| |||||| ||||  ||||       |  ||   |||| |    ||
Db        144 CAGGCACCTGTAATCCCAGGTACTAGGGAGGCTTTTGAACCCAGGAGGCAGAGGTTGCAG 85

Qy        265 ATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAA 324
                 || |||   ||   | |          |   |     ||||||||| :| ||||||||
Db         84 CGAGCTGAGATCGCGCCACTGCACTCCAGCCTGGGTGATAGAGTGAGATTCAGTCTCAAA 25

Qy        325 AACAACAACAAAAAACAAAAAAA 347
              || || || |||||| |||||||
Db         24 AAAAAAAAAAAAAAAAAAAAAAA 2
```

RESULT 7
US-11-266-748A-403127
; Sequence 403127, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 403127
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-403127

Query Match 37.0%; Score 147.8; DB 8; Length 1000;
Best Local Similarity 68.4%; Pred. No. 3.4e-22;
Matches 221; Conservative 1; Mismatches 93; Indels 8; Gaps 1;

```
Qy          33 ACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATC 92
               || ||| | |  || |   | |     ||  | ||||      |    | ||||||||||
Db         677 ACACAATTTTACATTAAGAAATACTGTGCAGGCCATGCGTGGTGGCTCAGGCCTGTAATC 736

Qy          93 CCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCC 152
               ||||||||| |||||||| |||||||| ||||||| ||| |||||||| ||
Db         737 CCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACCGGAGGTCAGGAGTTCGAGACCAGCC 796

Qy         153 TGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGG 212
               | |||||||| |||||||| |||||||||||||||||||||| |||| ||||||||||
Db         797 TTGCCAACATAGTGAAACCCTGTCTTTACTAAAAATACAAAAATTAGCCGGGCATGGTGG 856

Qy         213 CACACACCTGTAGTCCCAGCTACTCAGGAG--------CCGGAGATTGCAGTGAGCTGAG 264
               ||  |||||||| |||||| ||||  ||||          |  ||   |||| |     ||
Db         857 CAGGCACCTGTAATCCCAGGTACTAGGGAGGCTTTTGAACCCAGGAGGCAGAGGTTGCAG 916

Qy         265 ATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAA 324
               || |||   ||   | |          |   |      |||||||| :| ||||||||
Db         917 CGAGCTGAGATCGCGCCACTGCACTCCAGCCTGGGTGATAGAGTGAGATTCAGTCTCAAA 976

Qy         325 AACAACAACAAAAAACAAAAAAA 347
               || || || |||||| |||||||
Db         977 AAAAAAAAAAAAAAAAAAAAAAA 999
```

RESULT 8
US-11-266-748A-474173/c
; Sequence 474173, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 474173
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-474173

Query Match 37.0%; Score 147.8; DB 8; Length 1000;
Best Local Similarity 68.4%; Pred. No. 3.4e-22;
Matches 221; Conservative 1; Mismatches 93; Indels 8; Gaps 1;

```
Qy          33 ACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATC 92
               || ||| | |  || |   | |     ||  | ||||      |    | ||||||||||
Db         324 ACACAATTTTACATTAAGAAATACTGTGCAGGCCATGCGTGGTGGCTCAGGCCTGTAATC 265

Qy          93 CCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCC 152
               ||||||||| |||||||| |||||||| ||||||| ||| |||||||| ||
Db         264 CCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACCGGAGGTCAGGAGTTCGAGACCAGCC 205

Qy         153 TGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGG 212
               | |||||||| |||||||| |||||||||||||||||||||| |||| ||||||||||
Db         204 TTGCCAACATAGTGAAACCCTGTCTTTACTAAAAATACAAAAATTAGCCGGGCATGGTGG 145

Qy         213 CACACACCTGTAGTCCCAGCTACTCAGGAG--------CCGGAGATTGCAGTGAGCTGAG 264
               ||  |||||||| |||||| ||||  ||||          |  ||   |||| |     ||
Db         144 CAGGCACCTGTAATCCCAGGTACTAGGGAGGCTTTTGAACCCAGGAGGCAGAGGTTGCAG 85

Qy         265 ATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAA 324
               || |||   ||   | |          |   |      |||||||| :| ||||||||
Db          84 CGAGCTGAGATCGCGCCACTGCACTCCAGCCTGGGTGATAGAGTGAGATTCAGTCTCAAA 25

Qy         325 AACAACAACAAAAAACAAAAAAA 347
               || || || |||||| |||||||
Db          24 AAAAAAAAAAAAAAAAAAAAAAA 2
```

RESULT 9
US-11-319-952-56
; Sequence 56, Application US/11319952
; Publication No. US20060134114A1
; GENERAL INFORMATION:
; APPLICANT: Yousef, George M.
; APPLICANT: Diamandis, Eleftherios
; TITLE OF INVENTION: Novel Human Kallikrein-Like Genes
; FILE REFERENCE: MTS3USA
; CURRENT APPLICATION NUMBER: US/11/319,952
; CURRENT FILING DATE: 2005-12-28
; PRIOR APPLICATION NUMBER: US/09/936,271
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: PCT/CA00/00258

```
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: US 60/124,260
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/127,386
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: US 60/144,919
; PRIOR FILING DATE: 1999-07-21
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
;   LENGTH: 11820
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-319-952-56

  Query Match             36.9%;  Score 147.6;  DB 8;  Length 11820;
  Best Local Similarity   75.4%;  Pred. No. 4.2e-22;
  Matches  199;  Conservative    1;  Mismatches   55;  Indels    9;  Gaps    1;

Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
              |||||||||||| |||||| |||| ||||||| ||| | |  | ||||| || ||| ||
Db       2706 CCTGTAATCCCGGCACTTTGGGAAGCCAAGGCAGGCAGGTTGCTTGAGGCCAGGAGTTCA 2765

Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
              |||||| ||||||| |||||||||||||||||||| |||||||||||||||| |||| ||
Db       2766 AGACCAGCCTGGCCCACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCGG 2825

Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
              ||||||||||   |||||| |||||| |||||| |||  |||| |||||||||| ||
Db       2826 ACATGGTGGCACGTGCCTGTAATCCCAGGTACTCAAGAGGTGGAGGTTGCAGTGAGCGGA 2885

Qy        264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
              ||||  |||       |   | |   | | |         ||||| |||||:|| ||||||
Db       2886 GATCACAGCCCTGCACTCCAGCCTGGGTGA---------CAGAGCGAGACTCCATCTCAA 2936

Qy        324 AAACAACAACAAAAAACAAAAAAA 347
              ||  || ||||||||||| | ||
Db       2937 AACGAAAAACAAAAAACAGCACAA 2960


RESULT 10
US-11-266-748A-219455
; Sequence 219455, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 219455
;   LENGTH: 897
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-219455

  Query Match             36.8%;  Score 147.2;  DB 8;  Length 897;
  Best Local Similarity   75.4%;  Pred. No. 4.4e-22;
  Matches  196;  Conservative    1;  Mismatches   59;  Indels    4;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||| ||||||| |||||||| |||  || ||||||||||||||| |||
Db       593 ACGCCTGTAATCCTAGCACTTTGGGAGGCCGAGGCAGGTGGATCACCTGAGGTCAGGAGT 652

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| ||||||||||||||||||||| || ||| |||||||||||||| ||| ||||
Db       653 TCGAGACCAGCCTGGCCAACATGGTGAAACTCCATCTCTACTAAAAATACAGAAATTAGC 712

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||| |||||||  ||  |||||| |||||||||||  |||| | |||   | ||
Db       713 TGGGTATGGTGGTGCATGCCTGTAATCCCAGCTACTAGGGAGGCTGAGGCAGGAGAATCA 772

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              || | ||||  | |  |    ||  || |    ||    ||||||||||:| ||||
Db       773 CTTGAACCCAGAAGGTGAAGGGCACAGTGACC----AGGCGCCAGAGTGAGACTCTGTCT 828

Qy        321 CAAAAACAACAACAAAAAAC 340
              |||||| || || |||||||
Db       829 CAAAAAGAAAAAGAAAAAAC 848


RESULT 11
US-11-266-748A-239441/c
; Sequence 239441, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 239441
;   LENGTH: 897
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-239441

  Query Match             36.8%;  Score 147.2;  DB 8;  Length 897;
  Best Local Similarity   75.4%;  Pred. No. 4.4e-22;
  Matches  196;  Conservative    1;  Mismatches   59;  Indels    4;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||| ||||||| |||||||| |||  || ||||||||||||||| |||
Db       305 ACGCCTGTAATCCTAGCACTTTGGGAGGCCGAGGCAGGTGGATCACCTGAGGTCAGGAGT 246
```

```
Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||||||||||||||||||| || ||| |||||||||||||| ||| ||||
Db        245 TCGAGACCAGCCTGGCCAACATGGTGAAACTCCATCTCTACTAAAAATACAGAAATTAGC 186

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||| |||||||  ||  |||||| ||||||||||  |||| | |||   | ||
Db        185 TGGGTATGGTGGTGCATGCCTGTAATCCCAGCTACTAGGGAGGCTGAGGCAGGAGAATCA 126

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              || | ||||  | |  |    ||  || |    ||    ||||||||||:| ||||
Db        125 CTTGAACCCAGAAGGTGAAGGGCACAGTGACC----AGGCGCCAGAGTGAGACTCTGTCT 70

Qy        321 CAAAAACAACAACAAAAAAC 340
              ||||||  || || |||||||
Db         69 CAAAAAGAAAAAGAAAAAAC 50


RESULT 12
US-11-293-697-1113/c
; Sequence 1113, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/11/293,697
;  CURRENT FILING DATE:  2005-12-05
;  PRIOR APPLICATION NUMBER: US/10/108,260
;  PRIOR FILING DATE: 2002-03-28
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1113
;   LENGTH: 2909
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-293-697-1113

  Query Match             36.8%;  Score 147;  DB 8;  Length 2909;
  Best Local Similarity   71.9%;  Pred. No. 5.2e-22;
  Matches  192;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||| |||||||||||||| |||||||  |||  || |||||||||||||||| |||
Db        678 ACGCCTATAATCCCAGCACTTTGGGAGGCTGAGGCCGGTGGATCACCTGAGGTCAGGAGT 619

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || |||||| |||||||||||  |||||||||| ||| |||||||||||||||| ||||
Db        618 TCAAGACCAGCCTGGCCAACACAGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 559

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||||| |||  ||  |||||| ||||||||||  |||| | ||| ||||||||||
Db        558 TGGGCATGATGGTGCATGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGTTGCAGTGAGC 499

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
              ||||||||  |    |   | |   | | ||||||  ||       |
Db        498 CAAGATCGCACAACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCGTCTCCCAAAAAA 439

Qy        321 CAAAAACAACAACAAAAAACAAAAAAA 347
              ||| | |  || |||| | ||| | |
Db        438 TAAATAAATAAATAAAATAAAAATACA 412


RESULT 13
US-11-266-748A-23290/c
; Sequence 23290, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23290
;   LENGTH: 112595
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-23290

  Query Match             36.6%;  Score 146.2;  DB 8;  Length 112595;
  Best Local Similarity   78.5%;  Pred. No. 9.1e-22;
  Matches  175;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||||| |||||||| |||||||| |||||| ||| | || |||
Db      23489 ACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACTTGAAGCCAGGAGT 23430

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              |||||||| |||||||||||||||||||||||| |||||||||| |||| ||||
Db      23429 TCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACGAAAATTAGC 23370

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||||||||  ||| |||||||||||||||||| |||| |||||   | ||
Db      23369 CGGGCATGGTGGTGCACGCCTGTAGTCCCAGCTACTCGGGAGGCGGAGGCAGGAGAATCA 23310

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
              |  || ||| |  || |   |    ||  || |  ||| || ||
Db      23309 TTTGAACGCGGGTTGCGAAGGTTGCAGTGAGCCGAGATTGCGC 23267


RESULT 14
US-11-266-748A-30503/c
; Sequence 30503, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
```

```
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 30503
;   LENGTH: 6372
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-30503

  Query Match             36.4%;  Score 145.6;  DB 8;  Length 6372;
  Best Local Similarity   72.7%;  Pred. No. 1.1e-21;
  Matches  202;  Conservative    1;  Mismatches   70;  Indels    5;  Gaps    1;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||||| ||||||||||||  || |||| |||||||||| |||
Db       2899 ACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCAGGTGGATTACCTGAGGTCAGGAGT 2840

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
              || |||||| |||||||||||||||||||||||||||| |||||||||| ||||| |||
Db       2839 TCAAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATATAAAAATTAGT 2780

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGA-----GATTGCAG 255
              ||||| ||||||||     |||||| | |||||||||  || ||| ||     ||||||
Db       2779 TGGGCGTGGTGGCATGTGCCTGTAATGCCAGCTACTAGGGGGCCTGAGGCAGGATTGCTT 2720

Qy        256 TGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKC 315
              ||| ||| ||       ||    | |  |          |  ||  | |||||| ||||:|
Db       2719 TGACCTGGGAGGCAGAGGTTGCGCCACTTTCACTCCAGCCTAGGCAACAGAGTAAGACTC 2660

Qy        316 CGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAA 353
              |  |||||||| || || ||||||   ||||   | ||
Db       2659 CAGCTCAAAAAAAAAAAAAAAAAAGTAAAACGAACAA 2622


RESULT 15
US-11-266-748A-24156/c
; Sequence 24156, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24156
;   LENGTH: 3324
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-24156

  Query Match             36.3%;  Score 145;  DB 8;  Length 3324;
  Best Local Similarity   79.3%;  Pred. No. 1.4e-21;
  Matches  172;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||
Db        389 ATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGTGGATCACCTGAGGTCAAGAGT 330

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || |||||| |||||  ||||||||||||||||| ||| ||||||||||||||||| ||||
Db        329 TCAAGACCAGCCTGGGAAACATGGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 270

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||| ||||||| | |||||||| |||||||||||| |||| | ||| | | ||
Db        269 TGGGCGTGGTGGCGCGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGTAGGAGCATCA 210

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGA 297
                 || | |||   | |  |    ||  || |  |||
Db        209 CTTGAACCCAGGAGGTGGAGGCTGCAGTGAGCCGAGA 173


RESULT 16
US-11-266-748A-390100
; Sequence 390100, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 390100
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-390100

  Query Match             36.2%;  Score 144.8;  DB 8;  Length 1000;
  Best Local Similarity   68.1%;  Pred. No. 1.4e-21;
  Matches  218;  Conservative    1;  Mismatches   93;  Indels    8;  Gaps    1;

Qy         33 ACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATC 92
              || ||| | |  || |   | |      ||  | ||||       |    | ||||||||||
Db        434 ACACAATTTTACATTAAGAAATACTGTGCAGGCCATGCGTGGTGGCTCAGGCCTGTAATC 493

Qy         93 CCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCC 152
              ||||||||| |||||||| |||||||| ||||||| ||||||| ||| ||||||||| ||
Db        494 CCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACCGGAGGTCAGGAGTTCGAGACCAGCC 553

Qy        153 TGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGG 212
              | |||||||| ||||||||| |||| |||||||||||||||| |||| ||||||||||
```

```
Db        554 TTGCCAACATAGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCCGGGCATGGTGG 613

Qy        213 CACACACCTGTAGTCCCAGCTACTCAGGAG--------CCGGAGATTGCAGTGAGCTGAG 264

Db        614 CAGGCACCTGTAATCCCAGCTACTAGGGAGGCTTTTGAACCCAGGAGGCAGAGGTTGCAG 673

Qy        265 ATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAA 324

Db        674 CGAGCTGAGATCGCGCCACTGCACTCCAGCCTGGGTGATAGAGTGAGATTCAGTCTCAAA 733

Qy        325 AACAACAACAAAAAACAAAA 344

Db        734 AAAAAAAAAAAAAAAAAAAA 753


RESULT 17
US-11-266-748A-480818/c
; Sequence 480818, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 480818
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-480818

  Query Match             36.2%;  Score 144.8;  DB 8;  Length 1000;
  Best Local Similarity   68.1%;  Pred. No. 1.4e-21;
  Matches  218;  Conservative    1;  Mismatches   93;  Indels    8;  Gaps    1;

Qy         33 ACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCCTGTAATC 92

Db        567 ACACAATTTTACATTAAGAAATACTGTGCAGGCCATGCGTGGTGGCTCAGGCCTGTAATC 508

Qy         93 CCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAGACCATCC 152

Db        507 CCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACCGGAGGTCAGGAGTTCGAGACCAGCC 448

Qy        153 TGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGCATGGTGG 212

Db        447 TTGCCAACATAGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCCGGGCATGGTGG 388

Qy        213 CACACACCTGTAGTCCCAGCTACTCAGGAG--------CCGGAGATTGCAGTGAGCTGAG 264

Db        387 CAGGCACCTGTAATCCCAGCTACTAGGGAGGCTTTTGAACCCAGGAGGCAGAGGTTGCAG 328

Qy        265 ATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAAA 324

Db        327 CGAGCTGAGATCGCGCCACTGCACTCCAGCCTGGGTGATAGAGTGAGATTCAGTCTCAAA 268

Qy        325 AACAACAACAAAAAACAAAA 344

Db        267 AAAAAAAAAAAAAAAAAAAA 248


RESULT 18
US-11-266-748A-60135
; Sequence 60135, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 60135
;   LENGTH: 227968
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-60135

  Query Match             36.1%;  Score 144.4;  DB 8;  Length 227968;
  Best Local Similarity   76.1%;  Pred. No. 2.2e-21;
  Matches  178;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy         70 ACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTG 129

Db      40152 ACAGTGGCTCAACGCCTATAATCCCAGCACTTTGGGAGGCCGAGGTGGCCGGATCACCTG 40211

Qy        130 AGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATA 189

Db      40212 AGGTCAGCAGTTCAAGACCAGCCTGGCCAACGTGGTGAAACCCCGTCTCTACTAAAAACA 40271

Qy        190 CAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGA 249

Db      40272 CAAAAATTAGCTGGGCGTGGTGGTGTACACCTGTAGTCCCAGTTACTTAGGAGGCTGAGG 40331

Qy        250 TTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303

Db      40332 TAGGAGAATCAATTGAACCTGGGCAGCGGAGCTTGCAGTGAGCCGAGATAGAGC 40385


RESULT 19
US-11-266-748A-206136/c
; Sequence 206136, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
```

```
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 206136
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-206136

  Query Match             36.1%;  Score 144.2;  DB 8;  Length 1000;
  Best Local Similarity   70.0%;  Pred. No. 1.9e-21;
  Matches  194;  Conservative    0;  Mismatches   83;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||||||||||||||||| | || ||| ||| |||| ||||||||||||||| |||
Db         850 ACGCCTGTAATCCCAGCACTTTGAGAAGCCGAGGAGGGCAGATCACCTGAGGTCAGGAGT 791

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               || ||||||| ||| |||||||||||||||||||||   |||||||||||||||| ||||
Db         790 TCCAGACCAGCCTCGCCAACATGGTGAAACCCCGTACCTACTAAAAATACAAAAATTAGC 731

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||||||||||||||  |||||| || || |||||  |||| | ||||  | ||
Db         730 TGGGCATGGTGGCACATGCCTGTAATCTCACCTACTTGGGAGGCTGAGACAGGAGAATTG 671

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCT 320
                  || | ||||  | |  |    ||  || |  ||| || |||  |     |  | |
Db         670 CTTGAACCCAGAAGGTGAAGGTTGCAGTGAACCAAGATTGCGCACTGCACTCCAGCATGG 611

Qy         321 CAAAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
                  || |||  |||       |||||||  | || | |
Db         610 GCAAGACAGAAACTCTGTCAAAAAAAAAAAAAAAAA 574


RESULT 20
US-11-266-748A-60135/c
; Sequence 60135, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 60135
;   LENGTH: 227968
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-60135

  Query Match             36.1%;  Score 144.2;  DB 8;  Length 227968;
  Best Local Similarity   73.6%;  Pred. No. 2.5e-21;
  Matches  217;  Conservative    1;  Mismatches   59;  Indels   18;  Gaps    2;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | |||| ||||||||  | || ||||||||| | ||||| ||||||||||||| |||
Db      204862 ACGCCTATAATCCCATTAGTTTGGGAGGCCAACGCGGGCGTATCACCTGAGGTCAGGAGT 204803

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               ||||||||| ||||||||||||||||||||||| ||| ||||||||||||||||| ||||
Db      204802 TCGAGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGC 204743

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-------GCCGGAGATTGC 253
               |||| ||||||||||  |||||| ||||||||||||||||       || ||||| |
Db      204742 CGGGCGTGGTGGCACATGCCTGTAATCCCAGCTACTCAGGAGACTGAGGCAGGAGAATCG 204683

Qy         254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCA-----------CAGAGTGAG 302
                  ||| | | |   || |||   || |  | |  ||||||           |  ||  |
Db      204682 CTTGAACCGGGGAAGCGGAGGTTGCAGTGAGCCAAGATCAACCACTGCACTCCTAGGTAA 204623

Qy         303 CAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
               |||||  ||||:| |||||||||| || || |||||| |||||||  | || | |
Db      204622 CAGAGCTAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 204568


RESULT 21
US-11-266-748A-198112/c
; Sequence 198112, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
```

```
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 198112
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-198112

  Query Match           35.8%;  Score 143;  DB 8;  Length 1000;
  Best Local Similarity   77.6%;  Pred. No. 3.3e-21;
  Matches  173;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||| ||||||||||||||||  |||||||  |||||||  |||||||||||||||||  |||
Db        725 ATGCTTGTAATCCCAGCACTTTAGGAGGCCGAGGTGGGGGGATCACCTGAGGTCAGGAGG 666

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              |||||||||  |||| ||||||||||||||||||  |||| ||||||||||||||||||  ||||
Db        665 TCGAGACCAGCCTGCCCAACATGGTGAAACCCGGTCTCTACTAAAAATACAAAAATTAGC 606

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              ||||  |||||||||  ||||||||||||  ||||||||||||||  | |||    | ||
Db        605 TGGGAGTGGTGGCACGCACCTGTAGTCTCAGCTACTCAGGAGGCTGAGGCAGGAGAATCA 546

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
              || | |||   |     |    ||  || | ||||  | ||
Db        545 CTTGAACCCAGGAAGCAGAGGTTGCAGTGAGCCCAGATCGTGC 503


RESULT 22
US-11-266-748A-59101/c
; Sequence 59101, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 59101
;   LENGTH: 96217
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-59101

  Query Match           35.8%;  Score 143;  DB 8;  Length 96217;
  Best Local Similarity   88.6%;  Pred. No. 4.2e-21;
  Matches  155;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | ||||||||||||||||||  |||||||  |||||||  |||||||||||||||||  |||
Db      25922 ACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCACCTGAGGTCAGGAGT 25863

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              || ||||||  |||||||||||||||||||||||||||||  ||||||||||||||||  ||||
Db      25862 TCAAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 25803

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG 255
              ||||||||||||  |||  ||||||  ||||||||||||||||||  | |||    | ||
Db      25802 TGGGCATGGTGGTGCACGCCTGTAATCCCAGCTACTCAGGAGACTGAGGCAGGAG 25748


RESULT 23
US-11-266-748A-27021
; Sequence 27021, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 27021
;   LENGTH: 1945
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-27021

  Query Match           35.7%;  Score 142.6;  DB 8;  Length 1945;
  Best Local Similarity   69.4%;  Pred. No. 4.2e-21;
  Matches  209;  Conservative    1;  Mismatches   85;  Indels    6;  Gaps    1;

Qy         63 AGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGA 122
              ||||  | |    |  ||||||||  ||||||||||||  |||  |||  ||||||  ||||
Db       1636 AGGCCGGGCGTGTGGCTCATGCCTGTGATCCCAGCACTTTGGGTGGCTGAGGTGGACGGA 1695

Qy        123 TCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACT 182
              ||||||||||||  |||  | ||||||||||||||  |||  ||||||||||||||  ||||
Db       1696 TCACCTGAGGTCAGGAGTTTGAGACCATCCTGGCTAACGTGGTGAAACCCCGTCCCTACT 1755

Qy        183 AAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAG 242
              |||||||  ||||  ||||  ||||  |||  ||   | ||||||||||||||||||  ||||
Db       1756 AAAAATAACAAAATTAGCCGGGCGTGGGGGTGGGCGCCTGTAGTCCCAGCTACTCTGGAG 1815

Qy        243 CCGGAGATTGCAGTGAGCTGAGATC------GCAGAGTGAGCCGAAATCACAGATCACAG 296
              | |||   | ||   | | ||| |       |  |||   ||   | |   |
Db       1816 GCTGAGGCAGGAGAATGGCGTGAACCCGGGAGTCGAGATCGCGCCACTGCACTCCAGCCT 1875

Qy        297 AGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGAC 356
              |  | |||||  ||| :|||||||||||| || || ||||||  |||||||  | || |
Db       1876 GGGCAACAGAGCAAGAGTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1935

Qy        357 A 357
              |
```

```
Db       1936 A 1936


RESULT 24
US-11-266-748A-58369
; Sequence 58369, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 58369
;   LENGTH: 1945
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-58369

  Query Match             35.7%;  Score 142.6;  DB 8;  Length 1945;
  Best Local Similarity   69.4%;  Pred. No. 4.2e-21;
  Matches  209;  Conservative    1;  Mismatches   85;  Indels    6;  Gaps    1;

Qy         63 AGGCATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGA 122
              ||||  | |    |   |||||||| ||||||||||||| ||| |||  |||||| ||||
Db       1636 AGGCCGGGCGTGTGGCTCATGCCTGTGATCCCAGCACTTTGGGTGGCTGAGGTGGACGGA 1695

Qy        123 TCACCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACT 182
              ||||||||||||| ||| | |||||||||||||| ||| ||||||||||||||  ||||
Db       1696 TCACCTGAGGTCAGGAGTTTGAGACCATCCTGGCTAACGTGGTGAAACCCCGTCCCTACT 1755

Qy        183 AAAAATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAG 242
              |||||||  |||| |||| |||| ||| ||    | ||||||||||||||||||| ||||
Db       1756 AAAAATAACAAAATTAGCCGGGCGTGGGGGTGGGCGCCTGTAGTCCCAGCTACTCTGGAG 1815

Qy        243 CCGGAGATTGCAGTGAGCTGAGATC------GCAGAGTGAGCCGAAATCACAGATCACAG 296
              | |||   | ||   |  | || |       |  |||   ||   | |       |
Db       1816 GCTGAGGCAGGAGAATGGCGTGAACCCGGGAGTCGAGATCGCGCCACTGCACTCCAGCCT 1875

Qy        297 AGTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGAC 356
              |  | |||||  ||| :|||||||||||| || || |||||| |||||||  | || |
Db       1876 GGGCAACAGAGCAAGAGTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1935

Qy        357 A 357
              |
Db       1936 A 1936


RESULT 25
US-11-266-748A-28208
; Sequence 28208, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28208
;   LENGTH: 1421559
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (414394)..(414394)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (836909)..(836909)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1022504)..(1022504)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1041105)..(1041105)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1122667)..(1122667)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-28208

  Query Match             35.7%;  Score 142.6;  DB 8;  Length 1421559;
  Best Local Similarity   73.0%;  Pred. No. 5.8e-21;
  Matches  195;  Conservative    1;  Mismatches   70;  Indels    1;  Gaps    1;

Qy         85 CTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGA 144
              |||||||||| |||||| |||||||| ||||||| ||||||| ||||| || ||| ||||
Db    1069903 CTGTAATCCCGGCACTTTGGGAGGCCGAGGTGGGTGGATCACTTGAGGCCAGGAGTTCGA 1069962

Qy        145 GACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGG 204
              ||||| ||||||| |||||||||| ||||||| |||||||||||||||| ||||||||
Db    1069963 GACCAGCCTGGCCCACATGGTGAAAACCCGTCTCTACTAAAAATACAAAAATTAGCTGGG 1070022

Qy        205 CATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-GCCGGAGATTGCAGTGAGCTGA 263
              ||||||      |||||||||||||||| ||||| || | ||    |      |||
Db    1070023 TTTGGTGGTGTGTGCCTGTAGTCCCAGCTACCCAGGAGGCTGAAGCAAGAGAATCGCTTG 1070082

Qy        264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCCGTCTCAA 323
              |||  |||  | |           | || |    |||||||||| :| |||| |
Db    1070083 AATCTGGGAGGCGGAGGTTGCAGTCTTTAGCATGGGTGACAGAGTGAGATTCTGTCTTAG 1070142
```

```
Qy        324 AAACAACAACAAAAAACAAAAAAACCA 350
              |||||||||||| ||  | ||||||||
Db    1070143 AAACAACAACAACAACAAAAACCA 1070169


RESULT 26
US-11-266-748A-223929/c
; Sequence 223929, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 223929
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-223929

  Query Match             35.6%;  Score 142.4;  DB 8;  Length 1000;
  Best Local Similarity   69.9%;  Pred. No. 4.5e-21;
  Matches  207;  Conservative    1;  Mismatches   82;  Indels    6;  Gaps    1;

Qy         85 CTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGA 144
              |||||||||||||||| |||||||| |||  |  ||||||||||||||| ||| || |
Db        466 CTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGTTGGATCACCTGAGGTCAGGAGTTCAA 407

Qy        145 GACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGG 204
              ||||| |||||||||||| |||||||||||| ||| | ||||||||||||||| ||||||||
Db        406 GACCAGCCTGGCCAACACGGTGAAACCCCATCTCTGCTAAAAATACAAAAATTAGCTGGG 347

Qy        205 CATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAG 264
              | ||||||||  | |||||| |||||||||||| |||| | |||    | ||       |
Db        346 CGTGGTGGCAGGCGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTG 287

Qy        265 ATCGCAGAGTGAGCCGAAATCACAGATCA------CAGAGTGAGCAGAGTGAGACKCCGT 318
              | | |||   | |    ||  ||  ||      |   |    |||||||| ||:| ||
Db        286 AACCCAGGAGGCGAGATCATGCCATTGCACTCCAGCCTGGGTGACAGAGTGAAACTCTGT 227

Qy        319 CTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTC 374
              |||||||| || || |||| | || ||||        ||  || | |  |||   ||
Db        226 CTCAAAAAAAAAAAAAAAAGAAAAGAAAATTCCCTTACTGTGCCTACTTGCCCCTC 171


RESULT 27
US-11-266-748A-292231/c
; Sequence 292231, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 292231
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-292231

  Query Match             35.6%;  Score 142.4;  DB 8;  Length 1000;
  Best Local Similarity   69.9%;  Pred. No. 4.5e-21;
  Matches  207;  Conservative    1;  Mismatches   82;  Indels    6;  Gaps    1;

Qy         85 CTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGA 144
              |||||||||||||||| |||||||| |||  |  ||||||||||||||| ||| || |
Db        466 CTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGTTGGATCACCTGAGGTCAGGAGTTCAA 407

Qy        145 GACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGG 204
              ||||| |||||||||||| |||||||||||| ||| | ||||||||||||||| ||||||||
Db        406 GACCAGCCTGGCCAACACGGTGAAACCCCATCTCTGCTAAAAATACAAAAATTAGCTGGG 347

Qy        205 CATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAG 264
              | ||||||||  | |||||| |||||||||||| |||| | |||    | ||       |
Db        346 CGTGGTGGCAGGCGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTG 287

Qy        265 ATCGCAGAGTGAGCCGAAATCACAGATCA------CAGAGTGAGCAGAGTGAGACKCCGT 318
              | | |||   | |    ||  ||  ||      |   |    |||||||| ||:| ||
Db        286 AACCCAGGAGGCGAGATCATGCCATTGCACTCCAGCCTGGGTGACAGAGTGAAACTCTGT 227

Qy        319 CTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTC 374
              |||||||| || || |||| | || ||||        ||  || | |  |||   ||
Db        226 CTCAAAAAAAAAAAAAAAAGAAAAGAAAATTCCCTTACTGTGCCTACTTGCCCCTC 171


RESULT 28
US-11-266-748A-293957
; Sequence 293957, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
```

```
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 293957
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-293957

  Query Match             35.6%;  Score 142.4;  DB 8;  Length 1000;
  Best Local Similarity   73.0%;  Pred. No. 4.5e-21;
  Matches  219;  Conservative    1;  Mismatches   57;  Indels   23;  Gaps    2;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               ||||||||||||||||||||  |||||||| || ||| ||||| || ||||||| |||
Db         133 ATGCCTGTAATCCCAGCACTATGGGAGGCCAGGGAGGGAGGATCTCCAGAGGTCAGGAGT 192

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCC 200
               |||||||| |||||||||||||||||||||||||| || ||||||||||||| ||||
Db         193 TCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTATTAAAAATACAAAAATTAGC 252

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-------GCCGGAGATTGC 253
               |||||||| |||||  |||||||| |||||||||||| |||       || ||||| |
Db         253 TGGGCATGATGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATTG 312

Qy         254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCA----------------CAGA 297
               ||| ||| |   ||||||   || |   |    ||||||                |
Db         313 CTTGAACTGGGGAGGCAGAGGTTGCAGTGAGACAAGATCACGTCACTGTACTCCAGCCTG 372

Qy         298 GTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
               |    |||||||| ||:|| ||||||||| || || |||||| |||||||  | || | |
Db         373 GGTGACAGAGTGATACTCCATCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 432


RESULT 29
US-11-266-748A-343660
; Sequence 343660, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 343660
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-343660

  Query Match             35.6%;  Score 142.4;  DB 8;  Length 1000;
  Best Local Similarity   69.9%;  Pred. No. 4.5e-21;
  Matches  207;  Conservative    1;  Mismatches   82;  Indels    6;  Gaps    1;

Qy          85 CTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGA 144
               |||||||||||||||| |||||||| |||  |  |||||||||||||| ||| || |
Db         535 CTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGTTGGATCACCTGAGGTCAGGAGTTCAA 594

Qy         145 GACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGG 204
               ||||| |||||||||| |||||||||| ||| | |||||||||||||| ||||||||
Db         595 GACCAGCCTGGCCAACACGGTGAAACCCCATCTCTGCTAAAAATACAAAAATTAGCTGGG 654

Qy         205 CATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAG 264
               | ||||||||  | |||||| |||||||||||| |||| | |||   | ||       |
Db         655 CGTGGTGGCAGGCGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTG 714

Qy         265 ATCGCAGAGTGAGCCGAAATCACAGATCA------CAGAGTGAGCAGAGTGAGACKCCGT 318
               | | |||   | |     ||  ||   ||      |   |     |||||||| ||:| ||
Db         715 AACCCAGGAGGCGAGATCATGCCATTGCACTCCAGCCTGGGTGACAGAGTGAAACTCTGT 774

Qy         319 CTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTC 374
               |||||||| || || |||| | || ||||       ||  || | |  |||   ||
Db         775 CTCAAAAAAAAAAAAAAAAGAAAAGAAAATTCCCTTACTGTGCCTACTTGCCCCTC 830


RESULT 30
US-11-266-748A-345386/c
; Sequence 345386, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 345386
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-345386

  Query Match             35.6%;  Score 142.4;  DB 8;  Length 1000;
  Best Local Similarity   73.0%;  Pred. No. 4.5e-21;
```

```
  Matches  219;  Conservative    1;  Mismatches   57;  Indels   23;  Gaps    2;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               ||||||||||||||||||||  ||||||||| || ||| ||||| || ||||||| |||
Db         868 ATGCCTGTAATCCCAGCACTATGGGAGGCCAGGGAGGGAGGATCTCCAGAGGTCAGGAGT 809

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               ||||||||| ||||||||||||||||||||||||||| || ||||||||||||||| ||||
Db         808 TCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTATTAAAAATACAAAAATTAGC 749

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-------GCCGGAGATTGC 253
               |||||||| |||||  |||||||| |||||||||||| |||       || ||||| |
Db         748 TGGGCATGATGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATTG 689

Qy         254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCA----------------CAGA 297
                 ||| ||| |   ||||||   || |  |    ||||||                  |
Db         688 CTTGAACTGGGGAGGCAGAGGTTGCAGTGAGACAAGATCACGTCACTGTACTCCAGCCTG 629

Qy         298 GTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
               |    |||||||| ||:|| ||||||||| || || |||||| |||||||  | || | |
Db         628 GGTGACAGAGTGATACTCCATCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 569


RESULT 31
US-11-266-748A-403868/c
; Sequence 403868, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 403868
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-403868

  Query Match            35.6%;  Score 142.4;  DB 8;  Length 1000;
  Best Local Similarity  69.9%;  Pred. No. 4.5e-21;
  Matches  207;  Conservative    1;  Mismatches   82;  Indels    6;  Gaps    1;

Qy          85 CTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGA 144
               |||||||||||||||| |||||||| |||  |  |||||||||||||||| ||| || | |
Db         466 CTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGTTGGATCACCTGAGGTCAGGAGTTCAA 407

Qy         145 GACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGG 204
               ||||| |||||||||||| |||||||||||| ||| | ||||||||||||||||| ||||||||
Db         406 GACCAGCCTGGCCAACACGGTGAAACCCCATCTCTGCTAAAAATACAAAAATTAGCTGGG 347

Qy         205 CATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAG 264
               | ||||||||  | |||||| |||||||||||| |||| | |||   | ||         |
Db         346 CGTGGTGGCAGGCGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTG 287

Qy         265 ATCGCAGAGTGAGCCGAAATCACAGATCA------CAGAGTGAGCAGAGTGAGACKCCGT 318
               | | |||   | |      || ||  ||      |   |   |||||||| ||:| ||
Db         286 AACCCAGGAGGCGAGATCATGCCATTGCACTCCAGCCTGGGTGACAGAGTGAAACTCTGT 227

Qy         319 CTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTC 374
               |||||||| || || |||| | || ||||       ||  || | |  |||   ||
Db         226 CTCAAAAAAAAAAAAAAAAGAAAAGAAAATTCCCTTACTGTGCCTACTTGCCCCTC 171


RESULT 32
US-11-266-748A-406117
; Sequence 406117, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 406117
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-406117

  Query Match            35.6%;  Score 142.4;  DB 8;  Length 1000;
  Best Local Similarity  73.0%;  Pred. No. 4.5e-21;
  Matches  219;  Conservative    1;  Mismatches   57;  Indels   23;  Gaps    2;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||| ||||||||| || ||| ||||| || ||||||| |||
Db         133 ATGCCTGTAATCCCAGCACTATGGGAGGCCAGGGAGGGAGGATCTCCAGAGGTCAGGAGT 192

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               ||||||||| |||||||||||||||||||||||| || |||||||||||||| ||||
Db         193 TCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTATTAAAAATACAAAAATTAGC 252

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-------GCCGGAGATTGC 253
               |||||||| |||||  |||||||| |||||||||||| |||       || ||||| |
Db         253 TGGGCATGATGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATTG 312

Qy         254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCA----------------CAGA 297
                 ||| ||| |   ||||||   || |  |    ||||||                  |
Db         313 CTTGAACTGGGGAGGCAGAGGTTGCAGTGAGACAAGATCACGTCACTGTACTCCAGCCTG 372

Qy         298 GTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
               |    |||||||| ||:|| ||||||||| || || |||||| |||||||  | || | |
Db         373 GGTGACAGAGTGATACTCCATCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 432
```

```
RESULT 33
US-11-266-748A-474914
; Sequence 474914, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 474914
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-474914

  Query Match             35.6%;  Score 142.4;  DB 8;  Length 1000;
  Best Local Similarity   69.9%;  Pred. No. 4.5e-21;
  Matches  207;  Conservative    1;  Mismatches   82;  Indels    6;  Gaps    1;

Qy          85 CTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGA 144
               |||||||||||||||||| |||||||| |||   | ||||||||||||||||| ||| || |
Db         535 CTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGTTGGATCACCTGAGGTCAGGAGTTCAA 594

Qy         145 GACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGG 204
               ||||| |||||||||| |||||||||| ||| | |||||||||||||||| ||||||||
Db         595 GACCAGCCTGGCCAACACGGTGAAACCCCATCTCTGCTAAAAATACAAAAATTAGCTGGG 654

Qy         205 CATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGAG 264
               | ||||||||  | |||||| |||||||||||| |||| | |||   | ||       |
Db         655 CGTGGTGGCAGGCGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTG 714

Qy         265 ATCGCAGAGTGAGCCGAAATCACAGATCA------CAGAGTGAGCAGAGTGAGACKCCGT 318
               | | |||   | |     ||  ||   ||      |   |    |||||||| ||:| ||
Db         715 AACCCAGGAGGCGAGATCATGCCATTGCACTCCAGCCTGGGTGACAGAGTGAAACTCTGT 774

Qy         319 CTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCCATCTGCGGTTC 374
               |||||||| || || |||| | || ||||       ||  || | |  |||   ||
Db         775 CTCAAAAAAAAAAAAAAAAGAAAAGAAAATTCCCTTACTGTGCCTACTTGCCCCTC 830


RESULT 34
US-11-266-748A-477163/c
; Sequence 477163, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 477163
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-477163

  Query Match             35.6%;  Score 142.4;  DB 8;  Length 1000;
  Best Local Similarity   73.0%;  Pred. No. 4.5e-21;
  Matches  219;  Conservative    1;  Mismatches   57;  Indels   23;  Gaps    2;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||  ||||||||| || ||| ||||| || ||||||| |||
Db         868 ATGCCTGTAATCCCAGCACTATGGGAGGCCAGGGAGGGAGGATCTCCAGAGGTCAGGAGT 809

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               ||||||||| |||||||||||||||||||||||| || ||||||||||||| ||||
Db         808 TCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTATTAAAAATACAAAAATTAGC 749

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-------GCCGGAGATTGC 253
               |||||||| |||||  |||||||| |||||||||||| |||       || ||||| |
Db         748 TGGGCATGATGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATTG 689

Qy         254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCA---------------CAGA 297
                  ||| ||| |   ||||||   || |   |    ||||||                |
Db         688 CTTGAACTGGGGAGGCAGAGGTTGCAGTGAGACAAGATCACGTCACTGTACTCCAGCCTG 629

Qy         298 GTGAGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACA 357
               |    |||||||| ||:|| |||||||| || || |||||| |||||||  | || | |
Db         628 GGTGACAGAGTGATACTCCATCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 569


RESULT 35
US-11-191-644-1
; Sequence 1, Application US/11191644
; Publication No. US20060141529A1
; GENERAL INFORMATION:
;  APPLICANT: KOLESKE, ANTHONY JOHN
;  APPLICANT:  BOYLE, SCOTT NILE
;  APPLICANT:  SCHWEITZER, BARRY
;  APPLICANT:  MICHAUD, GREG
;  APPLICANT:  PREDKI, PAUL
;  TITLE OF INVENTION: COMPOSITIONS, KITS AND ASSAYS CONTAINING REAGENTS
;  TITLE OF INVENTION:  DIRECTED TO CORTACTIN AND AN ARG/ABL PROTEIN KINASE
;  FILE REFERENCE: INV-1007-UT
;  CURRENT APPLICATION NUMBER: US/11/191,644
;  CURRENT FILING DATE:  2005-07-27
;  PRIOR APPLICATION NUMBER: 60/591,827
;  PRIOR FILING DATE: 2004-07-27
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
;   LENGTH: 157866
```

```
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-11-191-644-1

  Query Match           35.6%;  Score 142.2;  DB 7;  Length 157866;
  Best Local Similarity   73.6%;  Pred. No. 6.3e-21;
  Matches  229;  Conservative    1;  Mismatches   59;  Indels   22;  Gaps    3;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||||||||| |||||||| |||  |||||||||  ||||||||||||
Db      62586 ATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGGCGGATCAC--GAGGTCAAGAGA 62643

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||| ||||
Db      62644 TTGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 62703

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-------GCCGGAGATTGC 253
              || |||||||||  ||| |||||||||||||||||||||||||       || ||||| |
Db      62704 TGTGCATGGTGGTGCACGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATCG 62763

Qy        254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGAT------------CACAGAGTG 300
              ||| |  ||   ||||||   || |  | |  ||||                ||||  ||
Db      62764 CTTGAACCCAGGAGGCAGAGGTTGCAGTGAGCTGAGATTGCACCACTGCACTCCAGACTG 62823

Qy        301 AGCAGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTG 360
              | ||||||||:||||||||||||  | |  || ||||||  || | ||  |||     |
Db      62824 CTGACAGTGAGACTCCGTCTCAAAAAACAAACAAACAAACAACAACAACAACAGAAGGAG 62883

Qy        361 TCCATCTGCGG 371
              |  |||  | |
Db      62884 ACTGTCTTCTG 62894


RESULT 36
US-11-293-697-23/c
; Sequence 23, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/11/293,697
;  CURRENT FILING DATE:  2005-12-05
;  PRIOR APPLICATION NUMBER: US/10/108,260
;  PRIOR FILING DATE: 2002-03-28
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
;   LENGTH: 2800
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-293-697-23

  Query Match           35.5%;  Score 142;  DB 8;  Length 2800;
  Best Local Similarity   72.3%;  Pred. No. 5.7e-21;
  Matches  198;  Conservative    1;  Mismatches   71;  Indels    4;  Gaps    1;

Qy         76 GCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCA 135
              | |  | ||||||||||||||||||| |||||||| |||  || ||||||||||||||||
Db        847 GTTTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGGTGGATCACCTGAGGTCA 788

Qy        136 AGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAA 195
              ||| ||||||||| ||||||||||||||||||||||||||||| ||||||||||||||||
Db        787 GGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAA 728

Qy        196 ATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG 255
              |||||||||| ||||||||||  |||||| |||||||||| | |||| | |||   | ||
Db        727 TTAGCTGGGCGTGGTGGCACATGCCTGTAATCCCAGCTACCCGGGAGGCTGAGGCAGGAG 668

Qy        256 TGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKC 315
                 || | | |   | | |  ||          |     | | ||  ||||:|
Db        667 AATCACTGGAACCCGGGAGGCGGAGGCCTC----CAGTCTAGTCAAAAAAAGCAAGACTC 612

Qy        316 CGTCTCAAAAACAACAACAAAAAACAAAAAAACC 349
              |||||||||||||||   | |    ||||||||
Db        611 AGTCTCAAAAACAACCGAACAGCCAAACAAAACC 578


RESULT 37
US-11-293-697-604/c
; Sequence 604, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/11/293,697
;  CURRENT FILING DATE:  2005-12-05
;  PRIOR APPLICATION NUMBER: US/10/108,260
;  PRIOR FILING DATE: 2002-03-28
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 604
;   LENGTH: 3252
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-293-697-604

  Query Match           35.5%;  Score 142;  DB 8;  Length 3252;
  Best Local Similarity   77.5%;  Pred. No. 5.7e-21;
  Matches  172;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | ||||||||||||||| |||| |||||||  |||  || |||||||  |||||||| |||
Db       2938 ACGCCTGTAATCCCAGTACTTTGGGAGGCTGAGGGAGGTGGATCACGTGAGGTCAGGAGT 2879

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              | ||||||| |||||||||||||||||||||||||||| |||||| || ||||||| ||||
Db       2878 TTGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAATAAAACAAAAATTAGC 2819

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||  |||||| ||   |||||| |||||||||||||  ||| | ||| ||||||||||||
Db       2818 CAGGCGTGGTGGTACGTGCCTGTAATCCCAGCTACTCGAGAGGCTGAGGTTGCAGTGAGC 2759

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAG 302
              |||||||           ||   | |   |   ||||||||||
Db       2758 TGAGATCATGCCACTGCCCTCCAGCGTGGGCGACAGAGTGAG 2717


RESULT 38
US-11-266-748A-202571
; Sequence 202571, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
```

```
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 202571
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-202571

  Query Match            35.4%;  Score 141.6;  DB 8;  Length 1000;
  Best Local Similarity   74.0%;  Pred. No. 6.5e-21;
  Matches  194;  Conservative    1;  Mismatches   60;  Indels    7;  Gaps    1;

Qy          84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
               |||||||||||||||||| |||||||||||||||||||||||||||||||   ||  ||
Db         383 CCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCTGGAATTCA 442

Qy         144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
               |||||||||||||||||| |||||||||||| ||| ||||||| ||||| |||| ||
Db         443 AGACCATCCTGGCCAACATGGCGAAACCCCGTCTCTACCAAAAATAGAAAAATTAGCCGG 502

Qy         204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAG-------CCGGAGATTGCAGT 256
               |||||||||||||    |||| ||||||||||  ||||       | ||||| |   |
Db         503 GCATGGTGGCACATGTTTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCACTT 562

Qy         257 GAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAGAGTGAGACKCC 316
               || ||  |   |  ||  || | || |        |   |    ||||||||||:
Db         563 GAACTTGGGAGGTGAAGGTTGCAGTAAGCCACTCCAGCCTGGGTGACAGAGTGAGACTTT 622

Qy         317 GTCTCAAAAACAACAACAAAAA 338
               ||||||||| || || |||||
Db         623 GTCTCAAAAAAAAAAAAAAAA 644


RESULT 39
US-11-266-748A-23004/c
; Sequence 23004, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23004
;   LENGTH: 216387
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-23004

  Query Match            35.4%;  Score 141.6;  DB 8;  Length 216387;
  Best Local Similarity   69.1%;  Pred. No. 8.5e-21;
  Matches  192;  Conservative    1;  Mismatches   85;  Indels    0;  Gaps    0;

Qy          66 CATGACACTGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCA 125
               || |  || |||      ||||| |||||||||||||| ||||||||||||||||| |||||
Db       87903 CAAGTAACAGGCCAGGCGCCTGAAATCCCAGCACTTTGGGAGGCCAAGGTGGGCAGATCA 87844

Qy         126 CCTGAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAA 185
               | |||||||| ||| | ||||||| |||||||||||||||||||||||||| |||||||
Db       87843 CTTGAGGTCAGGAGTTTGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAA 87784

Qy         186 AATACAAAAAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCG 245
               |||| ||||| |||  |||||||||||||||  |||||| |||||||||||||| ||| |
Db       87783 AATATAAAAATTAGTCGGGCATGGTGGCACATGCCTGTAATCCCAGCTACTCATGAGGCT 87724

Qy         246 GAGATTGCAGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGCAG 305
               |   | ||          || |   |  | | |     ||  || || ||| | ||
Db       87723 ACGGCAGGAGAATCACCTGAACCTGGGTGGTGGAGGCTGCAGTGAGCAGAGATCGTGCCA 87664

Qy         306 AGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAA 343
               |  : |    |||    | || |||||| |||
Db       87663 CTATGGGTGACAAAGCAAGGCTCAAAAAAAAAAAAAA 87626


RESULT 40
US-11-266-748A-201568
; Sequence 201568, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 201568
;   LENGTH: 632
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-201568

  Query Match            35.4%;  Score 141.4;  DB 8;  Length 632;
  Best Local Similarity   77.1%;  Pred. No. 7e-21;
  Matches  172;  Conservative    0;  Mismatches   51;  Indels    0;  Gaps    0;

Qy          26 CATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACACTGGCTGAATGCC 85
               || ||||    |  |||     | | | |  |||||||||| |       |    | |||
Db         311 CAGGGGAGGGAAGTTATGTGCCAAAAAGTACCAGGCCAGGCAGGCGCGATGGCTCACGCC 370

Qy          86 TGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCGAG 145
               | |||||||||||||| |||||||||||||||| || |||| |||||||| ||| || ||
Db         371 TATAATCCCAGCACTTTGGGAGGCCAAGGTGGGAGGGTCACTTGAGGTCAGGAGTTCAAG 430
```

```
Qy        146 ACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGGGC 205
              |||| |||||||||||||||||||||| | | |||||||||||||||||| ||||||||
Db        431 ACCAGCCTGGCCAACATGGTGAAACCCCATATCTACTAAAAATACAAAAATTAGCTGGGT 490

Qy        206 ATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAG 248
              |||·|||| |   |||||||||| |||||||||||||| | |||
Db        491 ATGATGGCTCGTGCCTGTAGTCCTAGCTACTCAGGAGGCTGAG 533


RESULT 41
US-11-266-748A-59943/c
; Sequence 59943, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 59943
;   LENGTH: 2445
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-59943

  Query Match             35.4%;  Score 141.4;  DB 8;  Length 2445;
  Best Local Similarity   88.0%;  Pred. No. 7.5e-21;
  Matches  154;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| |||||||||||||||| |||||||||||||||| |||
Db        344 AGGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGTGGATCACCTGAGGTCAGGAGT 285

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              |||||||| |||||||| |||||||||||||||||||||||||||||||||| || |
Db        284 TCGAGACCAGCCTGGCCAATATGGTGAAACCCCGTCTTTACTAAAAATACAAAAATTACC 225

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG 255
              ||||||||||  |||||||| |||||||| |||||| | |||   | ||
Db        224 CAGGCATGGTGGCGGACACCTGTAATCCCAGCTACCCAGGAGGCTGAGGCAGGAG 170


RESULT 42
US-10-517-441-32/c
; Sequence 32, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
;  APPLICANT: FOEKENS, John
;  APPLICANT:  HARBECK, Nadia
;  APPLICANT:  KOENIG, Thomas
;  APPLICANT:  MAIER, Sabine
;  APPLICANT:  MARTENS, John
;  APPLICANT:  MODEL, Fabian
;  APPLICANT:  NIMMRICH, Inko
;  APPLICANT:  RUJAN, Tamas
;  APPLICANT:  SCHMITT, Armin
;  APPLICANT:  SCHMITT, Manfred
;  APPLICANT:  LOOK, Maxime P.
;  APPLICANT:  MARX, Almuth
;  APPLICANT:  HOEFLER, Heinz
;  TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
;  TITLE OF INVENTION:  proliferative disorders
;  FILE REFERENCE: 47675-93
;  CURRENT APPLICATION NUMBER: US/10/517,441
;  CURRENT FILING DATE:  2004-12-11
;  PRIOR APPLICATION NUMBER: PCT/EP2003/010881
;  PRIOR FILING DATE: 2003-10-01
;  PRIOR APPLICATION NUMBER: DE 10317955.0
;  PRIOR FILING DATE: 2003-04-17
;  PRIOR APPLICATION NUMBER: DE 10300096.8
;  PRIOR FILING DATE: 2003-01-07
;  PRIOR APPLICATION NUMBER: DE 10245779.4
;  PRIOR FILING DATE: 2002-10-01
;  NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 32
;   LENGTH: 6521
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-10-517-441-32

  Query Match             35.4%;  Score 141.4;  DB 6;  Length 6521;
  Best Local Similarity   88.0%;  Pred. No. 7.9e-21;
  Matches  154;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||| |||| ||||||||| |||||||| ||||| || |||||||||||||| |||
Db       2279 ATGCCTCTAATACCAGCACTTTGGGAGGCCGAGGTGTGCAGATCACCTGAGGTCAGGAGT 2220

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||||||||||||||||||||||| |||||||||||||||| ||||
Db       2219 TCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 2160

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG 255
              ||||||||||||||  |||||||| || |||||||||||||| | |||   | ||
Db       2159 TGGGCATGGTGGCAGGCACCTGTAATCTCAGCTACTCAGGAGGCTGAGGCAGGAG 2105


RESULT 43
US-11-266-748A-60244
; Sequence 60244, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
```

```
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 60244
;   LENGTH: 86654
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-60244

  Query Match            35.4%;  Score 141.4;  DB 8;  Length 86654;
  Best Local Similarity   72.3%;  Pred. No. 9e-21;
  Matches  206;  Conservative    1;  Mismatches   57;  Indels   21;  Gaps    1;

Qy         84 CCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGATCG 143
              |||||||||||||||||| ||||||||||||  | ||||||||||||||||| ||| |||
Db      22962 CCTGTAATCCCAGCACTTTGGGAGGCCAAGGCAGACGGATCACCTGAGGTCAGGAGTTCG 23021

Qy        144 AGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGCTGG 203
              ||||||| ||||||||||||||  |||||||||||| |||||||||||||||||| |||||||
Db      23022 AGACCAGCCTGGCCAACATGGCTAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGG 23081

Qy        204 GCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGCTGA 263
              ||||| |||||||||||||||| ||||||||||  |||| | |||   | ||
Db      23082 GCATGATGGCACACACCTGTATTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCACTT 23141

Qy        264 GATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC---------- 303
              || |   |   | |  |    ||  || |  |||  | |
Db      23142 GAACCTGGGAGGTGGAGGTTGCAGTGAGCCAAGATCGCACCACTGCACTCCAGCTTGGGC 23201

Qy        304 -AGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAA 347
               ||||||||||:||||||||||| || || |||||| || ||||
Db      23202 AAGAGTGAGACTCCGTCTCAAAAAAAAAAAGAAAAAAAAACAAAA 23246


RESULT 44
US-11-266-748A-23291/c
; Sequence 23291, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23291
;   LENGTH: 113853
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-23291

  Query Match            35.4%;  Score 141.4;  DB 8;  Length 113853;
  Best Local Similarity   88.0%;  Pred. No. 9.1e-21;
  Matches  154;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              |||||||||||||||| ||| |||||||| ||||||| |||||||||||||||| |||
Db     108547 ATGCCTGTAATCCCAGCGCTTTGGGAGGCCGAGGTGGGTGGATCACCTGAGGTCAGGAGT 108488

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||| |||||| |||||||||||||||| |||||||||||||||  ||||
Db     108487 TCGAGACCAGCCTGACCAACAAGGTGAAACCCCGTCTCTACTAAAAATACAAAAGTTAGC 108428

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAG 255
              ||||||||||||||||  |||||||||||| ||||| |||| | ||||  | ||
Db     108427 TGGGCATGGTGGCACATGCCTGTAGTCCCAGATACTCGGGAGGCTGAGACAGGAG 108373


RESULT 45
US-11-266-748A-58517/c
; Sequence 58517, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan,.Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 58517
;   LENGTH: 154394
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-58517

  Query Match            35.4%;  Score 141.4;  DB 8;  Length 154394;
  Best Local Similarity   77.1%;  Pred. No. 9.2e-21;
  Matches  172;  Conservative    0;  Mismatches   51;  Indels    0;  Gaps    0;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||| ||||||| |||||| |||||||| |||  || ||||||||||||||| |||
Db     139360 ACGCCTATAATCCCGGCACTTTGGGAGGCCCAGGCAGGTGGATCACCTGAGGTCAGGAGT 139301

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||||||||||||||||||||||||||||| ||||||||| |||||||  |||
Db     139300 TCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAACACAAAAATAAGC 139241

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
              |||||  ||||||| |||||||||| |||||||||||||||| | |||   | ||
Db     139240 TGGGCGTGGTGGCGCACACCTGTAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATCA 139181

Qy        261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
              || |   ||  | || |    ||  || |  ||| || ||
Db     139180 CTTGAACATGGAAGGTGCAGGTTGCAGTGAGCTGAGATTGTGC 139138
```

```
RESULT 46
US-11-266-748A-22662/c
; Sequence 22662, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22662
;   LENGTH: 164429
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-22662

  Query Match             35.4%;  Score 141.4;  DB 8;  Length 164429;
  Best Local Similarity   77.1%;  Pred. No. 9.3e-21;
  Matches  172;  Conservative    0;  Mismatches   51;  Indels    0;  Gaps    0;

Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               |||||||||||||||||||||| |||||||  ||| |||| ||||||||||||||| |||
Db       27612 ATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCAGATCACCTGAGGTCAGGAGT 27553

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               ||||||||| |||||||||||||||||||||||||||| |||||||||||||||| ||||
Db       27552 TCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC 27493

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGCAGTGAGC 260
               ||||||||||||  |||||||| |||||||||||| |||| |  ||   | ||
Db       27492 CGGGCATGGTGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTAAGGCAGGAGAATTA 27433

Qy         261 TGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
                 || | |||   |    |    ||  || |  |||  | ||
Db       27432 CTTGAACCCAGGAGGCAGAGGTTGCAGTGAGCCAAGATCGTGC 27390


RESULT 47
US-11-266-748A-24937/c
; Sequence 24937, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24937
;   LENGTH: 86215
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-24937

  Query Match             35.3%;  Score 141.2;  DB 8;  Length 86215;
  Best Local Similarity   67.9%;  Pred. No. 9.9e-21;
  Matches  197;  Conservative    0;  Mismatches   93;  Indels    0;  Gaps    0;

Qy          14 CATGTGCTGGGCCATGGGAACCCAAATATTAATAAGACATTGTCAGGCCAGGCATGACAC 73
               ||| |  ||   | |   ||   ||  |    ||     |  || |  ||||||  | ||
Db       67383 CATTTATTGTTTCTTTTTAAATGAACAAAATTTAGAGAAAAGTTATTCCAGGCCAGGCAT 67324

Qy          74 TGGCTGAATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGT 133
               |      |||||||||||||||||||||||  ||||||  ||| ||| ||||||||||||||
Db       67323 GGTGATCATGCCTGTAATCCCAGCACTTTAGGAGGCTGAGGCGGGTGGATCACCTGAGGT 67264

Qy         134 CAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAA 193
               || ||| |||||||||| ||| |||||||||||| |||||  |||| |||||||||||||||
Db       67263 CAGGAGTTCGAGACCAGCCTAGCCAACATGGTGCAACCCTGTCTCTACTAAAAATACAAA 67204

Qy         194 AAATAGCTGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGAGCCGGAGATTGC 253
               || |||| ||||||||||     |   |||||||||||||||||||||||| | ||| | |
Db       67203 AATTAGCCGGGCATGGTGATGTAAGTCTGTAGTCCCAGCTACTCAGGAGGCTGAGGTAGG 67144

Qy         254 AGTGAGCTGAGATCGCAGAGTGAGCCGAAATCACAGATCACAGAGTGAGC 303
               |          || | |||   | | |    ||  || |  ||| || ||
Db       67143 ACAATCACTTGAACCCAGGAGGCGGAGGTTGCAGTGAGCTGAGATTGTGC 67094


RESULT 48
US-11-266-748A-119261
; Sequence 119261, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
```

```
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 119261
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-119261

  Query Match            35.3%;  Score 141;  DB 8;  Length 1000;
  Best Local Similarity  68.8%;  Pred. No. 8.7e-21;
  Matches  229;  Conservative    1;  Mismatches   86;  Indels   17;  Gaps    2;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| ||||||||  || ||| ||||  |||||||||| |||
Db        200 ACGCCTGTAATCCCAGCACTTTGGGAGGCCCTGGAGGGTGGATTGCCTGAGGTCAGGAGT 259

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||| | ||||||||||||||||||||  ||||||||||||||| ||||
Db        260 TCGAGACCAGCCTGCCTAACATGGTGAAACCCCGTCTCCACTAAAAATACAAAAATTAGC 319

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-GCCGGAGATTGCAGTGAG 259
              |||||  |||||||||||  |||||||||||||||| ||| ||| || |||||  |    |||
Db        320 TGGGCGTGGTGGCACACGCCTGTAGTCCCAGCTCCTCGGGAGGCAGGAGAATCGCTTGAA 379

Qy        260 CTGAGATCGCAGAGTGAGCCGAAATCACAGATCACA----------------GAGTGAGC 303
              |    |   | ||||   || |   | |  ||||| ||                    |   |
Db        380 CCCCGGAGGTAGAGGTTGCAGTGAGCCGAGATCGCAACACTACACTCCAGCCTGGGCGAC 439

Qy        304 AGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCC 363
              ||||  ||||:|| |||||||| || || |||||| ||||||  |    |  || |
Db        440 AGAGCAAGACTCCTTCTCAAAAAAAAAAAAAAAAAAAAAAAAAGGCCGGGCGCGGTGGCT 499

Qy        364 ATCTGCGGTTCCCAGACTATTGCAGGAGACCAA 396
                || | ||   |  |  | |  || || || |
Db        500 CACTCCTGTAATCCCAGCACTCTAGAAGGCCGA 532


RESULT 49
US-11-266-748A-161425/c
; Sequence 161425, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 161425
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-161425

  Query Match            35.3%;  Score 141;  DB 8;  Length 1000;
  Best Local Similarity  68.8%;  Pred. No. 8.7e-21;
  Matches  229;  Conservative    1;  Mismatches   86;  Indels   17;  Gaps    2;

Qy         81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
              | |||||||||||||||||| ||||||||  || ||| ||||  |||||||||| |||
Db        801 ACGCCTGTAATCCCAGCACTTTGGGAGGCCCTGGAGGGTGGATTGCCTGAGGTCAGGAGT 742

Qy        141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
              ||||||||| |||| | ||||||||||||||||||||  ||||||||||||||| ||||
Db        741 TCGAGACCAGCCTGCCTAACATGGTGAAACCCCGTCTCCACTAAAAATACAAAAATTAGC 682

Qy        201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-GCCGGAGATTGCAGTGAG 259
              |||||  |||||||||||  |||||||||||||||| ||| ||| || |||||  |    |||
Db        681 TGGGCGTGGTGGCACACGCCTGTAGTCCCAGCTCCTCGGGAGGCAGGAGAATCGCTTGAA 622

Qy        260 CTGAGATCGCAGAGTGAGCCGAAATCACAGATCACA----------------GAGTGAGC 303
              |    |   | ||||   || |   | |  ||||| ||                    |   |
Db        621 CCCCGGAGGTAGAGGTTGCAGTGAGCCGAGATCGCAACACTACACTCCAGCCTGGGCGAC 562

Qy        304 AGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCC 363
              ||||  ||||:|| |||||||| || || |||||| ||||||  |    |  || |
Db        561 AGAGCAAGACTCCTTCTCAAAAAAAAAAAAAAAAAAAAAAAAAGGCCGGGCGCGGTGGCT 502

Qy        364 ATCTGCGGTTCCCAGACTATTGCAGGAGACCAA 396
                || | ||   |  |  | |  || || || |
Db        501 CACTCCTGTAATCCCAGCACTCTAGAAGGCCGA 469


RESULT 50
US-11-266-748A-407602
; Sequence 407602, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 407602
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-407602

  Query Match            35.3%;  Score 141;  DB 8;  Length 1000;
  Best Local Similarity  68.8%;  Pred. No. 8.7e-21;
  Matches  229;  Conservative    1;  Mismatches   86;  Indels   17;  Gaps    2;
```

```
Qy          81 ATGCCTGTAATCCCAGCACTTCGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCAAGAGA 140
               | ||||||||||||||||||| ||||||||  || ||| ||||  |||||||||| |||
Db         200 ACGCCTGTAATCCCAGCACTTTGGGAGGCCCTGGAGGGTGGATTGCCTGAGGTCAGGAGT 259

Qy         141 TCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTTTACTAAAAATACAAAAAATAGC 200
               ||||||||| |||| | ||||||||||||||||||||  |||||||||||||||| ||||
Db         260 TCGAGACCAGCCTGCCTAACATGGTGAAACCCCGTCTCCACTAAAAATACAAAAATTAGC 319

Qy         201 TGGGCATGGTGGCACACACCTGTAGTCCCAGCTACTCAGGA-GCCGGAGATTGCAGTGAG 259
               |||||| |||||||||||| |||||||||||||||| ||| ||| || ||||| |    |||
Db         320 TGGGCGTGGTGGCACACGCCTGTAGTCCCAGCTCCTCGGGAGGCAGGAGAATCGCTTGAA 379

Qy         260 CTGAGATCGCAGAGTGAGCCGAAATCACAGATCACA----------------GAGTGAGC 303
               |   |    | ||||   || |  | |  ||||| ||                |    |
Db         380 CCCCGGAGGTAGAGGTTGCAGTGAGCCGAGATCGCAACACTACACTCCAGCCTGGGCGAC 439

Qy         304 AGAGTGAGACKCCGTCTCAAAAACAACAACAAAAAACAAAAAAACCATAAGACATTGTCC 363
               ||||  ||||:|| |||||||||| || || |||||| ||||||  |       |  || |
Db         440 AGAGCAAGACTCCTTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGGCCGGGCGCGGTGGCT 499

Qy         364 ATCTGCGGTTCCCAGACTATTGCAGGAGACCAA 396
                 || | ||    |  |  | |  || || || |
Db         500 CACTCCTGTAATCCCAGCACTCTAGAAGGCCGA 532
```

Search completed: July 17, 2006, 21:48:24
Job time : 416 secs